Access DB# 57203

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30_____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY	Type of Search	Vendors and cost where applicable
Searcher: D. Schreiber	NA Sequence (#) 1	STN _____
Searcher Phone #: 308-4292	AA Sequence (#) _____	Dialog _____
Searcher Location: CM1 12E18	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: 11/2	Bibliographic _____	Dr.Link _____
Date Completed: 11/5	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: 5	Fulltext _____	Sequence Systems Compugen
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: 10	Other _____	Other (specify) _____

PTO-1590 (1-01)

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:47:29 ; Search time 10414.5 Seconds
(without alignments)
12052.456 Million cell updates/sec

Title: US-09-778-516-1
Perfect score: 8115
Sequence: 1 gatgtacgggccagatatac.....gcgtgtcgctagcgacggtc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*
44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3535	43.6	3660	3	LBAGAL	M23530 L.delbrueck
	2	3497.2	43.1	5059	9	AX009488	AX009488 Sequence
	3	3386.6	41.7	5015	3	LBALACZBUL	M55068 L.delbrueck
	4	1336.4	16.5	3840	56	EVE132038	AJ132038 Expressio
	5	1240.4	15.3	3853	9	AR098190	AR098190 Sequence
c	6	1238.4	15.3	2093	3	LP1PLPREP	M31223 Plasmid pLP
c	7	1193.8	14.7	2140	2	C30ORFR	J03319 Plasmid pC3
	8	1082	13.3	3986	56	PCDNA3ZEO	X90639 Cloning vec

9	1068.6	13.2	4026	9	AR098191	AR098191 Sequence
10	988	12.2	4597	9	AX060344	AX060344 Sequence
11	973	12.0	4283	10	I63120	I63120 Sequence 3
12	973	12.0	4283	10	I85496	I85496 Sequence 1
13	968.6	11.9	4050	56	AF053407	AF053407 Expressio
14	968.6	11.9	4227	56	AF053408	AF053408 Expressio
15	968.6	11.9	4625	56	AF053409	AF053409 Expressio
16	968.2	11.9	4657	56	AF053406	AF053406 Expressio
17	959	11.8	5732	9	AX027904	AX027904 Sequence
18	959	11.8	5732	56	AX027910	AX027910 Sequence
19	935.6	11.5	5432	9	AX026821	AX026821 Sequence
20	935.2	11.5	5729	9	AX027908	AX027908 Sequence
21	935.2	11.5	5729	56	AX027914	AX027914 Sequence
22	928.6	11.4	6094	9	AX027906	AX027906 Sequence
23	928.6	11.4	6094	56	AX027912	AX027912 Sequence
24	921.8	11.4	6285	9	AR018923	AR018923 Sequence
25	921.8	11.4	6285	9	AR066551	AR066551 Sequence
26	921.8	11.4	6285	10	I72268	I72268 Sequence 49
27	921.8	11.4	6285	10	I76835	I76835 Sequence 49
28	909.2	11.2	6525	56	AF063848	AF063848 Cloning v
29	909.2	11.2	6631	56	AF063585	AF063585 Cloning v
30	909.2	11.2	6706	56	AF063850	AF063850 Cloning v
31	909.2	11.2	7758	56	AF063849	AF063849 Cloning v
32	908.2	11.2	4109	56	PSPORT1	U12390 Cloning vec
c 33	907	11.2	3018	56	AF324725	AF324725 Cloning v
c 34	907	11.2	3141	56	AF324726	AF324726 Cloning v
35	906.8	11.2	7511	56	CVU37023	U37023 Cloning vec
36	906.8	11.2	7511	56	CVU37068	U37068 Cloning vec
37	906.4	11.2	2000	10	I03343	I03343 Sequence 3
c 38	906.4	11.2	2661	56	SYN8KMRCG	M17626 Synthetic p
39	906.4	11.2	2689	56	AF016541	AF016541 Cloning v
40	906.4	11.2	2704	9	AX018504	AX018504 Sequence
41	906.4	11.2	2820	56	AF222336	AF222336 Cloning v
42	906.4	11.2	2860	56	AF017063	AF017063 Cloning v
c 43	906.4	11.2	2999	56	SYNCCDBA	L38498 Cloning vec
c 44	906.4	11.2	2999	56	SYNCCDBB	L38499 Cloning vec
c 45	906.4	11.2	3024	3	PK9CCDB	L27082 Plasmid KIL

ALIGNMENTS

```
RESULT   1
LBAGAL
LOCUS       LBAGAL       3660 bp    DNA             BCT       26-APR-1993
DEFINITION  L.delbrueckii bulgaricus beta-galactosidase gene, complete cds.
ACCESSION   M23530
VERSION     M23530.1  GI:149546
KEYWORDS    galactosidase.
SOURCE      L.delbrueckii bulgaricus (strain B131) DNA.
  ORGANISM  Lactobacillus delbrueckii
            Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
            Lactobacillus.
REFERENCE   1  (bases 1 to 3660)
  AUTHORS   Schmidt,B.F., Adams,R.M., Requadt,C., Power,S. and Mainzer,S.E.
  TITLE     Expression and nucleotide sequence of the Lactobacillus bulgaricus
            beta-galactosidase gene cloned in Escherichia coli
  JOURNAL   J. Bacteriol. 171, 625-635 (1989)
  MEDLINE   89123132
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by B.F.Schmidt, 29-MAR-1989.
FEATURES             Location/Qualifiers
     source          1. .3660
                     /organism="Lactobacillus delbrueckii"
                     /db_xref="taxon:1584"
     RBS             452. .458
                     /note="ribosomal binding site"
     CDS             468. .3491
                     /note="beta-galactosidase"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA25240.1"
                     /db_xref="GI:149547"
                     /translation="MSNKLVKEKRVDQADLAWLTDPEVYEVNTIPPHSDHESFQSQEE
                     LEEGKSSLVQSLDGDWLIDYAENGQGPVNFYAEDFDDSNFKSVKVPGNLELQGFGQPQ
                     YVNVQYPWDGSEEIFPPQIPSKNPLASYVRYFDLDEAFWDKEVSLKFDGAATAIYVWL
                     NGHFVGYGEDSFTPSEFMVTKFLKKENNRLAVALYKYSSASWLEDQDFWRMSGLFRSV
                     TLQAKPRLHLEDLKLTASLTDNYQKGKLEVEANIAYRLPNASFKLEVRDSEGDLVAEK
                     LGPIRSEQLEFTLADLPVAAWSAEKPNLYQVRLYLYQAGSLLEVSRQEVGFRNFELKD
                     GIMYLNGQRIVFKGANRHEFDSKLGRAITEEDMIWDIKTMKRSNINAVRCSHYPNQSL
                     FYRLCDKYGLYVIDEANLESHGTWEKVGGHEDPSFNVPGDDQHWLGASLSRVKNMMAR
                     DKNHASILIWSLGNESYAGTVFAQMADYVRKADPTRVQHYEGVTHNRKFDDATQIESR
                     MYAPAKVIEEYLTNKPAKPFISVEYAHAMGNSVGDLAAYTALEKYPHYQGGFIWDWID
                     QGLEKDGHLLYGGDFDDRPTDYEFCGNGLVFADRTESPKLANVKALYANLKLEVKDGQ
                     LFLKNDNLFTNSSSYYFLTSLLVDGKLTYQSRPLTFGLEPGESGTFALPWPEVADEKG
                     EVVYRVTAHLKEDLPWADEGFTVAEAEEVAQKLPEFKPEGRPDLVDSDYNLGLKGNNF
                     QILFSKVKGWPVSLKYAGREYLKRLPEFTFWRALTDNDRGAGYGYDLARWENAGKYAR
                     LKDISCEVKEDSVLVKTAFTLPVALKGDLTVTYEVDGRGKIAVTADFPGAEEAGLLPA
                     FGLNLALPKELTDYRYYGLGPNESYPDRLEGNYLGIYQGAVKKNFSPYRPQETGNRSK
                     VRWYQLFDEKGGLEFTANGADLNLSALPYSAAQIEAADHAFELTNNYTWVRALSAQMG
                     VGGDDSWGQKVHPEFCLDAQKARQLRLVIQPLLLK"
     misc_feature    3477. .3531
                     /note="region of dyad symmetry"
BASE COUNT      941 a    927 c    996 g    796 t
ORIGIN      1bp upsteam of HindIII site.
```

```
  Query Match           43.6%;  Score 3535;  DB 3;  Length 3660;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 3590;  Conservative    0;  Mismatches    0;  Indels   45;  Gaps    1;

Qy   4480 ttggcagccagtctccgggcaattaatgaacttggacatggttgacgacccggtctttgc 4539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11 TTGGCAGCCAGTCTCCGGGCAATTAATGAACTTGGACATGGTTGACGACCCGGTCTTTG- 69

Qy   4540 aagccgaattcgaccacactggcggccgttactagggtatcgatccgataaaaagttagg 4599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     70 ----------------------------------------CCGATAAAAAGTTAGG 85

Qy   4600 cgacggctttgccctggtgccagcagacggtaaggtctacgcgccatttgccggtactgt 4659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     86 CGACGGCTTTGCCCTGGTGCCAGCAGACGGTAAGGTCTACGCGCCATTTGCCGGTACTGT 145

Qy   4660 ccgccagctggccaagacccggcactcgatcgtcctggaaaatgaacatggggtcttggt 4719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    146 CCGCCAGCTGGCCAAGACCCGGCACTCGATCGTCCTGGAAAATGAACATGGGGTCTTGGT 205

Qy   4720 cttgattcaccttggcctgggcacggtcaaattaaacgggactggctttgtcagctatgt 4779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    206 CTTGATTCACCTTGGCCTGGGCACGGTCAAATTAAACGGGACTGGCTTTGTCAGCTATGT 265

Qy   4780 tgaagagggcagccaggtagaagccggccagcagatcctggaattctgggacccggcgat 4839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    266 TGAAGAGGGCAGCCAGGTAGAAGCCGGCCAGCAGATCCTGGAATTCTGGGACCCGGCGAT 325

Qy   4840 caagcaggccaagctggacgacacggtaatcgtgaccgtcatcaacagcgaaactttcac 4899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    326 CAAGCAGGCCAAGCTGGACGACACGGTAATCGTGACCGTCATCAACAGCGAAACTTTCAC 385

Qy   4900 aaatagccagatgctcttgccgatcggccacagcgtccaagccctggatgatgtattcaa 4959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386 AAATAGCCAGATGCTCTTGCCGATCGGCCACAGCGTCCAAGCCCTGGATGATGTATTCAA 445

Qy   4960 gttagaagggaagaattagaaaatgagcaataagttagtaaaagaaaaaagagttgacca 5019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    446 GTTAGAAGGGAAGAATTAGAAAATGAGCAATAAGTTAGTAAAAGAAAAAAGAGTTGACCA 505

Qy   5020 ggcagacctggcctggctgactgacccggaagtttacgaagtcaatacaattcccccgca 5079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506 GGCAGACCTGGCCTGGCTGACTGACCCGGAAGTTTACGAAGTCAATACAATTCCCCCGCA 565

Qy   5080 ctccgaccatgagtccttccaaagccaggaagaactggaggagggcaagtccagtttagt 5139
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    566 CTCCGACCATGAGTCCTTCCAAAGCCAGGAAGAACTGGAGGAGGGCAAGTCCAGTTTAGT 625

Qy   5140 gcagtccctggacggggactggctgattgactacgctgaaaacggccagggaccagtcaa 5199
```

```
Db    626 GCAGTCCCTGGACGGGGACTGGCTGATTGACTACGCTGAAAACGGCCAGGGACCAGTCAA 685

Qy   5200 cttctatgcagaagactttgacgatagcaatttttaagtcagtcaaagtacccggcaacct 5259
Db    686 CTTCTATGCAGAAGACTTTGACGATAGCAATTTTAAGTCAGTCAAAGTACCCGGCAACCT 745

Qy   5260 ggaactgcaaggctttggccagccccagtatgtcaacgtccaatatccatgggacggcag 5319
Db    746 GGAACTGCAAGGCTTTGGCCAGCCCCAGTATGTCAACGTCCAATATCCATGGGACGGCAG 805

Qy   5320 tgaggagattttcccgccccaaattccaagcaaaaatccgctcgcttcttatgtcagata 5379
Db    806 TGAGGAGATTTTCCCGCCCCAAATTCCAAGCAAAAATCCGCTCGCTTCTTATGTCAGATA 865

Qy   5380 ctttgacctggatgaagctttctgggacaaggaagtcagcttgaagtttgacggggcggc 5439
Db    866 CTTTGACCTGGATGAAGCTTTCTGGGACAAGGAAGTCAGCTTGAAGTTTGACGGGGCGGC 925

Qy   5440 aacagccatctatgtctggctgaacggccacttcgtcggctacggggaagactcctttac 5499
Db    926 AACAGCCATCTATGTCTGGCTGAACGGCCACTTCGTCGGCTACGGGGAAGACTCCTTTAC 985

Qy   5500 cccaagcgagtttatggttaccaagttcctcaagaaagaaaataaccgcctggcagtggc 5559
Db    986 CCCAAGCGAGTTTATGGTTACCAAGTTCCTCAAGAAAGAAAATAACCGCCTGGCAGTGGC 1045

Qy   5560 tctctacaagtattcttccgcctcctggctggaagaccaggacttctggcgcatgtctgg 5619
Db   1046 TCTCTACAAGTATTCTTCCGCCTCCTGGCTGGAAGACCAGGACTTCTGGCGCATGTCTGG 1105

Qy   5620 tttgttcagatcagtgactcttcaggccaagccgcgtctgcacttggaggaccttaagct 5679
Db   1106 TTTGTTCAGATCAGTGACTCTTCAGGCCAAGCCGCGTCTGCACTTGGAGGACCTTAAGCT 1165

Qy   5680 tacggccagcttgaccgataactaccaaaaaggaaagctggaagtcgaagccaatattgc 5739
Db   1166 TACGGCCAGCTTGACCGATAACTACCAAAAAGGAAAGCTGGAAGTCGAAGCCAATATTGC 1225

Qy   5740 ctaccgcttgccaaatgccagctttaagctggaagtgcgggatagtgaaggtgacttggt 5799
Db   1226 CTACCGCTTGCCAAATGCCAGCTTTAAGCTGGAAGTGCGGGATAGTGAAGGTGACTTGGT 1285

Qy   5800 tgctgaaaagctgggcccaatcagaagcgagcagctggaattcactctggctgatttgcc 5859
Db   1286 TGCTGAAAAGCTGGGCCCAATCAGAAGCGAGCAGCTGGAATTCACTCTGGCTGATTTGCC 1345

Qy   5860 agtagctgcctggagcgcggaaaagcctaacctttaccaggtccgcctgtatttatacca 5919
Db   1346 AGTAGCTGCCTGGAGCGCGGAAAAGCCTAACCTTTACCAGGTCCGCCTGTATTTATACCA 1405

Qy   5920 ggcaggcagcctcttagaggttagccggcaggaagtgggtttccgcaactttgaactaaa 5979
Db   1406 GGCAGGCAGCCTCTTAGAGGTTAGCCGGCAGGAAGTGGGTTTCCGCAACTTTGAACTAAA 1465

Qy   5980 agacgggattatgtaccttaacggccagcggatcgtcttcaaggggccaaccggcacga 6039
Db   1466 AGACGGGATTATGTACCTTAACGGCCAGCGGATCGTCTTCAAGGGGCCAACCGGCACGA 1525

Qy   6040 atttgacagtaagttgggtcgggctatcacggaagaggatatgatctgggacatcaagac 6099
Db   1526 ATTTGACAGTAAGTTGGGTCGGGCTATCACGGAAGAGGATATGATCTGGGACATCAAGAC 1585

Qy   6100 catgaagcgaagcaacatcaatgctgtccgctgctctcactacccgaaccagtccctctt 6159
Db   1586 CATGAAGCGAAGCAACATCAATGCTGTCCGCTGCTCTCACTACCCGAACCAGTCCCTCTT 1645

Qy   6160 ttaccggctctgtgacaagtacggcctttacgtcattgatgaagctaacctggaaagcca 6219
Db   1646 TTACCGGCTCTGTGACAAGTACGGCCTTTACGTCATTGATGAAGCTAACCTGGAAAGCCA 1705

Qy   6220 cggcacctgggaaaaagtgggggggcacgaagatcctagcttcaatgttccaggcgatga 6279
Db   1706 CGGCACCTGGGAAAAAGTGGGGGGGCACGAAGATCCTAGCTTCAATGTTCCAGGCGATGA 1765

Qy   6280 ccagcattggctgggagccagcttatcccgggtgaagaacatgatggctcgggacaagaa 6339
Db   1766 CCAGCATTGGCTGGGAGCCAGCTTATCCCGGGTGAAGAACATGATGGCTCGGGACAAGAA 1825

Qy   6340 ccatgcttcaatcctaatctggtctttaggcaatgagtcttacgccggcactgtctttgc 6399
Db   1826 CCATGCTTCAATCCTAATCTGGTCTTTAGGCAATGAGTCTTACGCCGGCACTGTCTTTGC 1885

Qy   6400 ccaaatggctgattacgtccggaaggctgatccgacccgggttcagcactatgaaggggt 6459
Db   1886 CCAAATGGCTGATTACGTCCGGAAGGCTGATCCGACCCGGGTTCAGCACTATGAAGGGGT 1945

Qy   6460 gacccacaaccggaagtttgacgacgccacccagattgaaagccggatgtatgctccggc 6519
Db   1946 GACCCACAACCGGAAGTTTGACGACGCCACCCAGATTGAAAGCCGGATGTATGCTCCGGC 2005

Qy   6520 caaggtaattgaagaatacttgaccaataaaccagccaagccatttatctcagttgaata 6579
Db   2006 CAAGGTAATTGAAGAATACTTGACCAATAAACCAGCCAAGCCATTTATCTCAGTTGAATA 2065

Qy   6580 cgctcacgccatgggcaactccgtcggtgacctggccgcctacacggccctggaaaaata 6639
Db   2066 CGCTCACGCCATGGGCAACTCCGTCGGTGACCTGGCCGCCTACACGGCCCTGGAAAAATA 2125

Qy   6640 cccccactaccagggcggcttcatctgggactggattgaccaaggactggaaaaagacgg 6699
Db   2126 CCCCCACTACCAGGGCGGCTTCATCTGGGACTGGATTGACCAAGGACTGGAAAAAGACGG 2185

Qy   6700 gcacctgctttatgggggcgacttcgatgaccggccaaccgactatgaattctgcgggaa 6759
Db   2186 GCACCTGCTTTATGGGGGCGACTTCGATGACCGGCCAACCGACTATGAATTCTGCGGGAA 2245

Qy   6760 cggcctggtctttgctgaccggactgaatcgccgaaactggctaatgtcaaggccccttta 6819
Db   2246 CGGCCTGGTCTTTGCTGACCGGACTGAATCGCCGAAACTGGCTAATGTCAAGGCCCTTTA 2305

Qy   6820 cgccaaccttaagttagaagtaaaagatgggcagctcttcctcaaaaacgacaatttatt 6879
Db   2306 CGCCAACCTTAAGTTAGAAGTAAAAGATGGGCAGCTCTTCCTCAAAAACGACAATTTATT 2365

Qy   6880 taccaacagctcatcttactacttcttgactagtcttttggtcgatggcaagttgaccta 6939
Db   2366 TACCAACAGCTCATCTTACTACTTCTTGACTAGTCTTTTGGTCGATGGCAAGTTGACCTA 2425

Qy   6940 ccagagccggcctctgaccttggcctggagcctggcgaatccgggaccttgccctgcc 6999
Db   2426 CCAGAGCCGGCCTCTGACCTTTGGCCTGGAGCCTGGCGAATCCGGGACCTTTGCCCTGCC 2485

Qy   7000 ttggccggaagtcgctgatgaaaaaggggaggtcgtctaccgggtaacggcccacttaaa 7059
Db   2486 TTGGCCGGAAGTCGCTGATGAAAAAGGGGAGGTCGTCTACCGGGTAACGGCCCACTTAAA 2545

Qy   7060 agaagacttgccttgggcggatgagggcttcactgtggctgaagcagaagaagtagctca 7119
Db   2546 AGAAGACTTGCCTTGGGCGGATGAGGGCTTCACTGTGGCTGAAGCAGAAGAAGTAGCTCA 2605

Qy   7120 aaagctgccggaatttaagccggaagggcggccagatttagttgattccgactacaacct 7179
Db   2606 AAAGCTGCCGGAATTTAAGCCGGAAGGGCGGCCAGATTTAGTTGATTCCGACTACAACCT 2665

Qy   7180 aggcctgaaaggaaataacttccaaattctcttctccaaggtcaagggctggccggtttc 7239
Db   2666 AGGCCTGAAAGGAAATAACTTCCAAATTCTCTTCTCCAAGGTCAAGGGCTGGCCGGTTTC 2725

Qy   7240 cctcaagtatgccggtagggaatacttgaagcggctgccggaatttaccttctggcgggc 7299
Db   2726 CCTCAAGTATGCCGGTAGGGAATACTTGAAGCGGCTGCCGGAATTTACCTTCTGGCGGGC 2785

Qy   7300 cctgacggacaacgaccggggagctggttacggctatgatctggcccggtgggaaaatgc 7359
Db   2786 CCTGACGGACAACGACCGGGGAGCTGGTTACGGCTATGATCTGGCCCGGTGGGAAAATGC 2845
```

```
Qy   7360 cggcaagtatgcccgcttgaaagacatcagctgcgaggtcaaggaagactccgttttggt 7419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 CGGCAAGTATGCCCGCTTGAAAGACATCAGCTGCGAGGTCAAGGAAGACTCCGTTTTGGT 2905

Qy   7420 caagactgcctttacgttgcctgtcgccttaaagggtgatttaaccgtgacctatgaagt 7479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 CAAGACTGCCTTTACGTTGCCTGTCGCCTTAAAGGGTGATTTAACCGTGACCTATGAAGT 2965

Qy   7480 cgatggacggggcaagattgctgtaacagctgacttcccaggcgcggaagaagctggtct 7539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 CGATGGACGGGGCAAGATTGCTGTAACAGCTGACTTCCCAGGCGCGGAAGAAGCTGGTCT 3025

Qy   7540 cttgccagcctttggcttgaacctggccctgccaaaagaactgaccgattaccgctacta 7599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 CTTGCCAGCCTTTGGCTTGAACCTGGCCCTGCCAAAAGAACTGACCGATTACCGCTACTA 3085

Qy   7600 tggtctgggacctaatgagagctacccagaccgcttggaaggtaattacctgggcatcta 7659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 TGGTCTGGGACCTAATGAGAGCTACCCAGACCGCTTGGAAGGTAATTACCTGGGCATCTA 3145

Qy   7660 ccagggagcggtaaaaaagaactttagcccatatcgtccgcaggaaacgggcaaccggag 7719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 CCAGGGAGCGGTAAAAAAGAACTTTAGCCCATATCGTCCGCAGGAAACGGGCAACCGGAG 3205

Qy   7720 caaggttcgctggtaccagctctttgatgaaaagggcggcttggaatttacggccaatgg 7779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3206 CAAGGTTCGCTGGTACCAGCTCTTTGATGAAAAGGGCGGCTTGGAATTTACGGCCAATGG 3265

Qy   7780 ggcagacttgaacttgtctgctttgccatattctgccgcccaaattgaagcagcggacca 7839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 GGCAGACTTGAACTTGTCTGCTTTGCCATATTCTGCCGCCCAAATTGAAGCAGCGGACCA 3325

Qy   7840 cgcttttgaactgactaacaattacacttgggttagagccttaagcgcccagatgggggt 7899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3326 CGCTTTTGAACTGACTAACAATTACACTTGGGTTAGAGCCTTAAGCGCCCAGATGGGGGT 3385

Qy   7900 cggcggggatgactcctgggggcagaaggtccacccggaattctgcctggatgctcaaaa 7959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3386 CGGCGGGGATGACTCCTGGGGGCAGAAGGTCCACCCGGAATTCTGCCTGGATGCTCAAAA 3445

Qy   7960 agcccgccagcttcgcctggtgattcagcccctttactaaaataaatgctacaattgac 8019
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3446 AGCCCGCCAGCTTCGCCTGGTGATTCAGCCCCTTTTACTAAAATAAATGCTACAATTGAC 3505

Qy   8020 ttaacaggatgaaatttagtaaaagcaaagcgagtgaggaagatggcaacgatcagaga 8079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3506 TTAACAGGATGAAATTTTAGTAAAAGCAAAGCGAGTGAGGAAGATGGCAACGATCAGAGA 3565

Qy   8080 agtgccaaggcagccggcgtgtcgctagcgacggt 8114
          |||||||||||||||||||||||||||||||||
Db   3566 AGTGCCAAGGCAGCCGGCGTGTCGCTAGCGACGGT 3600
```

```
RESULT   2
AX009488
LOCUS       AX009488     5059 bp    DNA             PAT       06-SEP-2000
DEFINITION  Sequence 1 from Patent WO9961627.
ACCESSION   AX009488
VERSION     AX009488.1  GI:9996771
KEYWORDS    .
SOURCE      Lactobacillus delbrueckii subsp. bulgaricus.
  ORGANISM  Lactobacillus delbrueckii subsp. bulgaricus
            Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
            Lactobacillus.
REFERENCE   1  (bases 1 to 5059)
  AUTHORS   Brignon,P., Gendre,F. and Benbadis,L.
  TITLE     Mutant lactobacillus bulgaricus strains free from beta-galactoside
            activity
  JOURNAL   Patent: WO 9961627-A 1 02-DEC-1999;
            GERVAIS DANONE CO (FR); BRIGNON PIERRE (FR); GENDRE FRANCOIS (FR);
            BENBADIS LAURENT (FR)
FEATURES             Location/Qualifiers
     source          1. .5059
                     /organism="Lactobacillus delbrueckii subsp. bulgaricus"
                     /db_xref="taxon:1585"
     CDS             122. .1873
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC07487.1"
                     /db_xref="GI:9996772"
                     /translation="MIFIITNLITAIRIGEVLLDPLIGNAIDRTESRWGKFKPWVVGG
                     GIISSLALLALFTDFGGINQSNPVVYLVIFGIVYLIMDIFYSFKDTGFWAMIPALSLD
                     SREREKTSTFARVGSTIGANLVGVVITPIILFFSASKANPNGDKQGWFFFALIVAIVG
                     ILTSITVGLGTHEVKSALRESNEKTTLKQVFKVLGQNDQLLWLAFAYWFYGLGINTLN
                     ALQLYYFSYILGDARGYSLLYTINTFVGLISASFFPSLAKKFNRNRLFYACIAVMLLG
                     IGVFSVASGSLALSLVGAEFFFIPQPLAFLVVLMIISDAVEYGQLKTGHRDEALTLSV
                     RPLVDKLGGALSNWFVSLIALTAGMTTGATASTITAHGQMVFKLAMFALPAVMLLIAV
                     SIFAKKVFLTEEKHAEIVDQLETQFSQSHAQKPAQAESFTLASPVSGQLMNLDMVDDP
                     VFADKKLGDGFALVPADGKVYAPFAGTVRQLAKTRHSIVLENEHGVLVLIHLGLGTVK
                     LNGTGFVSYVEEGSQVEAGQQILEFWDPAIKQAKLDDTVIVTVINSETFANSQMLLPI
                     GHSVQALDDVFKLEGKN"
     CDS             1877. .4519
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC07488.1"
                     /db_xref="GI:9996773"
                     /translation="MSNKLVKEKRVDQADLAWLTDPEVYEVNTIPPHSDHESFQSQEE
                     LEEGKSSLVQSLDGDWLIDYAENGQGPVNFYAEDFDDSNFKSVKVPGNLELQGFGQPQ
                     YVNVQYPWDGSEEIFPPQIPSKNPLASYVRYFDLDEAFWDKEVSLKFDGAATAIYVWL
                     NGHFVGYGEDSFTPSEFMVTKFLKKENNRLAVALYKYSSASWLEDQDFWRMSGLFRSV
                     TLQAKPRLHLEDLKLTASLTDNYQKGKLEVEANIAYRLPNASFKLEVRDSEGDLVAEK
                     LGPIRSEQLEFTLADLPVAAWSAEKPNLYQVRLYLYQAGSLLEVSRQEVGFRNFELKD
                     GIMYLNGQRIVFKGANRHEFDSKLGRAITEEDMIWDIKTMKRSNINAVRCSHYPNQSL
                     FYRLCDKYGLYVIDEANLESHGTWEKVGGHEDPSFNVPGDDQHWLGASLSRVKNMMAR
                     DKNHASILIWSLGNESYAGTVFAQMADYVRKADPTRVQHYEGVTHNRKFDDATQIESR
                     MYAPAKVIEEYLTNKPAKPFISVEYAHAMGNSVGDLAAYTALEKYPHYQGGFIWDWID
                     QGLEKDGHLLYGGDFDDRPTDYEFCGNGLVFADRTESPKLANVKALYANLKLEVKDGQ
                     LFLKNDNLFTNSSSYYFLTSLLVDGKLTYQSRPLTFGLEPGESGTFALPWPEVADEKG
                     EVVYRVTAHLKEDLPWADEGFTVAEAEEVAQKLPEFKPEGRPDLVDSDYNLGLKGNNF
                     QILFSKVKGWPVSLKYAGREYLKRLPEFTFWRALTDNDRGAGYGYDLARWENAGKYAR
                     LKDISCEVKEDSVLVKTAFTLPVALKGDLTVTYEVDGRGKIAVTADFPGAEEAGLLPA
                     FGLNLALPKELTDYRYYGLGPNES"
BASE COUNT     1229 a   1341 c   1323 g   1166 t
ORIGIN
```

```
  Query Match             43.1%;  Score 3497.2;  DB 9;  Length 5059;
  Best Local Similarity   99.5%;  Pred. No. 0;
  Matches 3518;  Conservative    0;  Mismatches   13;  Indels    3;  Gaps    1;

Qy   4584 ccgataaaaagttaggcgacggctttgccctggtgccagcagacggtaaggtctacgcgc 4643
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1479 CCGACAAAAAGTTAGGCGACGGCTTTGCCCTGGTGCCAGCAGACGGTAAGGTCTACGCGC 1538

Qy   4644 catttgccggtactgtccgccagctggccaagacccggcactcgatcgtcctggaaaatg 4703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1539 CATTTGCCGGTACTGTCCGCCAGCTGGCCAAGACCCGGCACTCGATCGTCCTGGAAAATG 1598

Qy   4704 aacatggggtcttggtcttgattcaccttggcctgggcacggtcaaattaaacgggactg 4763
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1599 AACATGGGGTCTTGGTCTTGATTCACCTTGGCCTGGGCACGGTCAAATTAAACGGGACTG 1658

Qy   4764 gctttgtcagctatgttgaagagggcagccaggtagaagccggccagcagatcctggaat 4823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1659 GCTTTGTCAGCTATGTTGAAGAGGGCAGCCAGGTAGAAGCCGGCCAGCAGATCCTGGAAT 1718

Qy   4824 tctgggacccggcgatcaagcaggccaagctggacgacacggtaatcgtgaccgtcatca 4883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1719 TCTGGGACCCGGCGATCAAGCAGGCCAAGCTGGACGACACGGTAATCGTGACCGTCATCA 1778

Qy   4884 acagcgaaactttcacaaatagccagatgctcttgccgatcggccacagcgtccaagccc 4943
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1779 ACAGCGAAACTTTCGCAAATAGCCAGATGCTCTTGCCGATCGGCCACAGCGTCCAAGCCC 1838
```

```
Qy   4944 tggatgatgtattcaagttagaagggaagaattagaaaatgagcaataagttagtaaaag 5003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1839 TGGATGATGTATTCAAGTTAGAAGGGAAGAATTAGAAAATGAGCAATAAGTTAGTAAAAG 1898

Qy   5004 aaaaaagagttgaccaggcagacctggcctggctgactgacccggaagtttacgaagtca 5063
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1899 AAAAAAGAGTTGACCAGGCAGACTTGGCCTGGCTGACTGACCCGGAAGTTTACGAAGTCA 1958

Qy   5064 atacaattcccccgcactccgaccatgagtccttccaaagccaggaagaactggaggagg 5123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1959 ATACAATTCCCCCGCACTCCGACCATGAGTCCTTCCAAAGCCAGGAAGAACTGGAGGAGG 2018

Qy   5124 gcaagtccagtttagtgcagtccctggacggggactggctgattgactacgctgaaaacg 5183
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2019 GCAAGTCCAGTTTAGTGCAGTCCCTGGACGGGGACTGGCTGATTGACTACGCTGAAAACG 2078

Qy   5184 gccagggaccagtcaacttctatgcagaagactttgacgatagcaattttaagtcagtca 5243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2079 GCCAGGGACCAGTCAACTTCTATGCAGAAGACTTTGACGATAGCAATTTTAAGTCAGTCA 2138

Qy   5244 aagtacccggcaacctggaactgcaaggctttggccagccccagtatgtcaacgtccaat 5303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2139 AAGTACCCGGCAACCTGGAACTGCAAGGCTTTGGCCAGCCCCAGTATGTCAACGTCCAAT 2198

Qy   5304 atccatgggacggcagtgaggagattttcccgccccaaattccaagcaaaaatccgctcg 5363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2199 ATCCATGGGACGGCAGTGAGGAGATTTTCCCGCCCCAAATTCCAAGCAAAAATCCGCTCG 2258

Qy   5364 cttcttatgtcagatactttgacctggatgaagctttctgggacaaggaagtcagcttga 5423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2259 CTTCTTATGTCAGATACTTTGACCTGGATGAAGCTTTCTGGGACAAGGAAGTCAGCTTGA 2318

Qy   5424 agtttgacggggcggcaacagccatctatgtctggctgaacggccacttcgtcggctacg 5483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2319 AGTTTGACGGGGCGGCAACAGCCATCTATGTCTGGCTGAACGGCCACTTCGTCGGCTACG 2378

Qy   5484 gggaagactcctttaccccaagcgagtttatggttaccaagttcctcaagaaagaaaata 5543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2379 GGGAAGACTCCTTTACCCCAAGCGAGTTTATGGTTACCAAGTTCCTCAAGAAAGAAAATA 2438

Qy   5544 accgcctggcagtggctctctacaagtattcttccgcctcctggctggaagaccaggact 5603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2439 ACCGCCTGGCAGTGGCTCTCTACAAGTATTCTTCCGCCTCCTGGCTGGAAGACCAGGACT 2498

Qy   5604 tctggcgcatgtctggtttgttcagatcagtgactcttcaggccaagccgcgtctgcact 5663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2499 TCTGGCGCATGTCTGGTTTGTTCAGATCAGTGACTCTTCAGGCCAAGCCGCGTCTGCACT 2558

Qy   5664 tggaggaccttaagcttacggccagcttgaccgataactaccaaaaaggaaagctggaag 5723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2559 TGGAGGACCTTAAGCTTACGGCCAGCTTGACCGATAACTACCAAAAAGGAAAGCTGGAAG 2618

Qy   5724 tcgaagccaatattgcctaccgcttgccaaatgccagctttaagctggaagtgcgggata 5783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2619 TCGAAGCCAATATTGCCTACCGCTTGCCAAATGCCAGCTTTAAGCTGGAAGTGCGGGATA 2678

Qy   5784 gtgaaggtgacttggttgctgaaaagctgggcccaatcagaagcgagcagctggaattca 5843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2679 GTGAAGGTGACTTGGTTGCTGAAAAGCTGGGCCCAATCAGAAGCGAGCAGCTGGAATTCA 2738

Qy   5844 ctctggctgatttgccagtagctgcctggagcgcggaaaagcctaacctttaccaggtcc 5903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2739 CTCTGGCTGATTTGCCAGTAGCTGCCTGGAGCGCGGAAAAGCCTAACCTTTACCAGGTCC 2798

Qy   5904 gcctgtatttataccaggcaggcagcctcttagaggttagccggcaggaagtgggtttcc 5963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2799 GCCTGTATTTATACCAGGCAGGCAGCCTCTTAGAGGTTAGCCGGCAGGAAGTGGGTTTCC 2858

Qy   5964 gcaactttgaactaaaagacgggattatgtaccttaacggccagcggatcgtcttcaagg 6023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2859 GCAACTTTGAACTAAAAGACGGGATTATGTACCTTAACGGCCAGCGGATCGTCTTCAAGG 2918

Qy   6024 gggccaaccggcacgaatttgacagtaagttgggtcgggctatcacggaagaggatatga 6083
          |||||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||
Db   2919 GGGCCAACCGGCACGAATTTGACAGTAAGTTGGGCCGGGCTATCACAGAAGAGGATATGA 2978

Qy   6084 tctgggacatcaagaccatgaagcgaagcaacatcaatgctgtccgctgctctcactacc 6143
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2979 TCTGGGATATCAAGACCATGAAGCGAAGCAACATCAATGCTGTCCGCTGCTCTCACTACC 3038

Qy   6144 cgaaccagtccctcttttaccggctctgtgacaagtacggcctttacgtcattgatgaag 6203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3039 CGAACCAGTCCCTCTTTTACCGGCTCTGTGACAAGTACGGCCTTTACGTCATTGATGAAG 3098

Qy   6204 ctaacctggaaagccacggcacctgggaaaaagtgggggggcacgaagatcctagcttca 6263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3099 CTAACCTGGAAAGCCACGGCACCTGGGAAAAAGTGGGGGGGCACGAAGATCCTAGCTTCA 3158

Qy   6264 atgttccaggcgatgaccagcattggctgggagccagcttatcccgggtgaagaacatga 6323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 ATGTTCCAGGCGATGACCAGCATTGGCTGGGAGCCAGCTTATCCCGGGTGAAGAACATGA 3218

Qy   6324 tggctcgggacaagaaccatgcttcaatcctaatctggtctttaggcaatgagtcttacg 6383
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3219 TGGCTCGGGACAAGAACCATGCTTCAATCCTGATCTGGTCTTTAGGCAATGAGTCTTACG 3278

Qy   6384 ccggcactgtctttgcccaaatggctgattacgtccggaaggctgatccgacccgggttc 6443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 CCGGCACTGTCTTTGCCCAAATGGCTGATTACGTCCGGAAGGCTGATCCGACCCGGGTTC 3338

Qy   6444 agcactatgaaggggtgacccacaaccggaagtttgacgacgccacccagattgaaagcc 6503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3339 AGCACTATGAAGGGGTGACCCACAACCGGAAGTTTGACGACGCCACCCAGATTGAAAGCC 3398

Qy   6504 ggatgtatgctccggccaaggtaattgaagaatacttgaccaataaaccagccaagccat 6563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3399 GGATGTATGCTCCGGCCAAGGTAATTGAAGAATACTTGACCAATAAACCAGCCAAGCCAT 3458

Qy   6564 ttatctcagttgaatacgctcacgccatgggcaactccgtcggtgacctggccgcctaca 6623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3459 TTATCTCAGTTGAATACGCTCACGCCATGGGCAACTCCGTCGGTGACCTGGCCGCCTACA 3518

Qy   6624 cggccctggaaaaatacccccactaccagggcggcttcatctgggactggattgaccaag 6683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3519 CGGCCCTGGAAAAATACCCCCACTACCAGGGCGGCTTCATCTGGGACTGGATTGACCAAG 3578

Qy   6684 gactggaaaaagacgggcacctgctttatgggggcgacttcgatgaccggccaaccgact 6743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3579 GACTGGAAAAAGACGGGCACCTGCTTTATGGGGGCGACTTCGATGACCGGCCAACCGACT 3638

Qy   6744 atgaattctgcgggaacggcctggtctttgctgaccggactgaatcgccgaaactggcta 6803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3639 ATGAATTCTGCGGGAACGGCCTGGTCTTTGCTGACCGGACTGAATCGCCGAAACTGGCTA 3698

Qy   6804 atgtcaaggccctttacgccaaccttaagttagaagtaaaagatgggcagctcttcctca 6863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3699 ATGTCAAGGCCCTTTACGCCAACCTTAAGTTAGAAGTAAAAGATGGGCAGCTCTTCCTCA 3758

Qy   6864 aaaacgacaatttatttaccaacagctcatcttactacttcttgactagtcttttggtcg 6923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3759 AAAACGACAATTTATTTACCAACAGCTCATCTTACTACTTCTTGACTAGTCTTTTGGTCG 3818

Qy   6924 atggcaagttgacctaccagagccggcctctgacctttggcctggagcctggcgaatccg 6983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3819 ATGGCAAGTTGACCTACCAGAGCCGGCCTCTGACCTTTGGCCTGGAGCCTGGCGAATCCG 3878

Qy   6984 ggacctttgccctgccttggccggaagtcgctgatgaaaaaggggaggtcgtctaccggg 7043
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   3879 GGACCTTTGCCCTGCCTTGGCCGGAAGTCGCTGATGAAAAAGGAGAGGTCGTCTACCGGG 3938

Qy   7044 taacggcccacttaaaagaagacttgccttgggcggatgagggcttcactgtggctgaag 7103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3939 TAACGGCCCACTTAAAAGAAGACTTGCCTTGGGCGGATGAGGGCTTCACTGTGGCTGAAG 3998

Qy   7104 cagaagaagtagctcaaaagctgccggaatttaagccggaagggcggccagatttagttg 7163
```

```
Db    3999 CAGAAGAAGTAGCTCAAAAGCTGCCGGAATTTAAGCCGGAAGGGCGGCCAGATTTAGTTG 4058

Qy    7164 attccgactacaacctaggcctgaaaggaaataacttccaaattctcttctccaaggtca 7223
Db    4059 ATTCCGACTACAACCTAGGCCTGAAAGGAAATAACTTCCAAATTCTCTTCTCCAAGGTCA 4118

Qy    7224 agggctggccggtttccctcaagtatgccggtagggaatacttgaagcggctgccggaat 7283
Db    4119 AGGGCTGGCCGGTTTCCCTCAAGTATGCCGGTAGGGAATACTTGAAGCGGCTGCCGGAAT 4178

Qy    7284 ttaccttctggcgggccctgacggacaacgaccggggagctggttacggctatgatctgg 7343
Db    4179 TTACCTTCTGGCGGGCCCTGACGGACAACGACCGGGGAGCTGGTTACGGCTATGATCTGG 4238

Qy    7344 cccggtgggaaaatgccggcaagtatgcccgcttgaaagacatcagctgcgaggtcaagg 7403
Db    4239 CCCGGTGGGAAAATGCCGGCAAGTATGCCCGCTTGAAAGACATCAGCTGCGAGGTCAAGG 4298

Qy    7404 aagactccgttttggtcaagactgcctttacgttgcctgtcgccttaaagggtgatttaa 7463
Db    4299 AAGACTCCGTTTTGGTCAAGACTGCCTTTACGTTGCCTGTCGCCTTAAAGGGTGATTTAA 4358

Qy    7464 ccgtgacctatgaagtcgatggacggggcaagattgctgtaacagctgacttcccaggcg 7523
Db    4359 CTGTGACCTATGAAGTCGATGGACGGGGCAAGATTGCTGTAACAGCTGACTTCCCAGGCG 4418

Qy    7524 cggaagaagctggtctcttgccagcctttggcttgaacctggccctgccaaaagaactga 7583
Db    4419 CGGAAGAAGCCGGTCTCTTGCCAGCCTTTGGCTTGAACCTGGCCCTGCCAAAAGAACTGA 4478

Qy    7584 ccgattaccgctactatggtctgggacctaatgagagctacccagaccgcttggaaggta 7643
Db    4479 CCGATTACCGCTACTATGGTCTGGGACCTAATGAGAGCTAACCAGACCGCTTGGAAGGTA 4538

Qy.   7644 attacctgggcatctaccagggagcggtaaaaaagaactttagcccata---tcgtccgc 7700
Db    4539 ATTACCTGGGCATCTACCAGGGAGCGGTAAAAAAGAACTTTAGCCCATACCTGCGTCCGC 4598

Qy    7701 aggaaacgggcaaccggagcaaggttcgctggtaccagctctttgatgaaaagggcggct 7760
Db    4599 AGGAAACGGGCAACCGGAGCAAGGTTCGCTGGTACCAGCTCTTTGATGAAAAGGGCGGCT 4658

Qy    7761 tggaatttacggccaatggggcagacttgaacttgtctgctttgccatattctgccgccc 7820
Db    4659 TGGAATTTACGGCCAATGGGGCAGACTTGAACTTGTCTGCTTTGCCATATTCTGCCGCCC 4718

Qy    7821 aaattgaagcagcggaccacgcttttgaactgactaacaattacacttgggttagagcct 7880
Db    4719 AAATTGAAGCAGCGGACCACGCTTTTGAACTGACTAACAATTACACTTGGGTTAGAGCCT 4778

Qy    7881 taagcgcccagatgggggtcggcggggatgactcctgggggcagaaggtccacccggaat 7940
Db    4779 TAAGCGCCCAGATGGGGGTCGGCGGGGATGACTCCTGGGGGCAGAAGGTCCACCCGGAAT 4838

Qy    7941 tctgcctggatgctcaaaaagcccgccagcttcgcctggtgattcagccccttttactaa 8000
Db    4839 TCTGCCTGGATGCTCAAAAAGCCCGCCAGCTCCGCCTGGTGATTCAGCCCCTTTTACTAA 4898

Qy    8001 aataaatgctacaattgacttaacaggatgaaattttagtaaaagcaaagcgagtgagga 8060
Db    4899 AATAAATGCTACAATTGACTTAACAGGATGAAATTTTAGTAAAAGCAAAGCGAGTGAGGA 4958

Qy    8061 agatggcaacgatcagagaagtgccaaggcagccggcgtgtcgctagcgacggt 8114
Db    4959 AGATGGCAACGATCAGAGAAGTGCCAAGGCAGCCGGCGTGTCGCTAGCGACGGT 5012


RESULT   3
LBALACZBUL
LOCUS       LBALACZBUL   5015 bp    DNA             BCT       26-APR-1993
DEFINITION  L.delbrueckii bulgaricus lactose permease and beta-galactosidase
            (lacZ) genes, complete cds.
            "
ACCESSION   M55068 M38754
VERSION     M55068.1  GI:149564
KEYWORDS    beta-galactosidase; lactose permease.
SOURCE      L.delbrueckii bulgaricus DNA.
  ORGANISM  Lactobacillus delbrueckii
            Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
            Lactobacillus.
REFERENCE   1  (bases 1 to 5015)
  AUTHORS   Leong-Morgenthaler,P.M., Zwahlen,M.C. and Hottinger,H.
  TITLE     Lactose metabolism in Lactobacillus bulgaricus: Analysis of the
            primary structure and expression of the genes involved
  JOURNAL   J. Bacteriol. 173, 1951-1957 (1991)
  MEDLINE   91161509
FEATURES             Location/Qualifiers
     source          1. .5015
                     /organism="Lactobacillus delbrueckii"
                     /sub_species="bulgaricus"
                     /db_xref="taxon:1584"
                     /tissue_lib="ATCC 11842"
     gene            102. .1985
                     /gene="lactose permease"
     CDS             102. .1985
                     /gene="lactose permease"
                     /note="putative"
                     /codon_start=1
                     /transl_table=11
                     /product="lactose permease"
                     /protein_id="AAA25243.1"
                     /db_xref="GI:149565"
                     /translation="MKKKLVSRLSYAAGAFGNDVFYATLSTYFIVFVTTHLFNAGDHK
                     MIFIITNLITAIRIGEVLLDPLIGNAIDRTESRWGKFKPWVVGGGIISSLALLALFTD
                     FGGINQSKPVVYLVIFGIVYLIMDIFYSFKDTGFWAMIPALSLDSRERERKTSTFARVG
                     STIGANLVGVVITPIILFFSASKANPNGDKQGWFFFALIVAIVGILTSITVGLGTHEV
                     KSALRESNEKTTLKQVFKVLGQNDQLLWLAFAYWFYGLGINTLNALQLYYFSYILGDA
                     RGYSLLYTINTFVGLISASFFPSLAKKFNRNRLFYACIAVMLLGIGVFSVASGSLALS
                     LVGAEFFFIPQPLAFLVVLMIISDAVEYGQLKTGHRDEALTLSVRPLVDKLGGALSNW
                     FVSLIALTAGMTTGATASTITAHGQMVFKLAMFALPAVMLLIAVSIFAKKVFLTEEKH
                     AEIVDQLETQFGQSHAQKPAQAESFTLASPVSGQLMNLDMVDDPVFADKKLGDGFALV
                     PADGKVYAPFAGTVRQLAKTRHSIVLENEHGVLVLIHLGLGTAKLNGTGFVSYVEEGS
                     QVEAGQQILEFWDPAIKQAKLDDTVIVTVINSETFANSQMLLPIGHSVQALDDVFKLE
                     GKN"
     gene            1989. .5015
                     /gene="lacZ"
     CDS             1989. .5015
                     /gene="lacZ"
                     /note="putative"
                     /codon_start=1
                     /transl_table=11
                     /product="beta-galactosidase"
                     /protein_id="AAA25244.1"
                     /db_xref="GI:149566"
                     /translation="MSNKLVKEKRVDQADLAWLTDPEVYEVNTIPPHSDHESFQSQEE
                     LEEGKSSLVQSLDGDWLIDYAENGQGPVNFYAEDFDDSNFKSVKVPGNLELQGFGQPQ
                     YVNVQYPWDGSEEIFPPQIPSKNPLASYVRYFDLDEAFWDKEVSLKFDGAAATAIYVWL
                     NGHFVGYGEDSFTPSEFMVTKFLKKENNRLAVALYKYSSASWLEDQDFWRMSGLFRSV
                     TLQAKPRLHLEDLKLTASLTDNYQKGKLEVEANIAYRLPNASFKLEVRDSEGDLVAEK
                     LGPIRSEQLEFTLADLPVAAWSAEKPNLYQVRLYLYQAGSLLEVSRQEVGFRNFELKD
                     GIMYLNGQRIVFKGANRHEFDSKLGRAITEEDMIWDIKTMKRSNINAVRCSHYPNQSL
                     FYRLCDKYGLYVIDEANLESHGTWEKVGGHEDPSFNVPGDDQHWLGASLSRVKNMMAR
                     DKNHASILIWSLGNESYAGTVFAQMADYVRKADPTRVQHYEGVTHNRKFDDATQIESR
                     MYAPAKVIEEYLTNKPAKPFISVEYAHAMGNSVGDLAAYTALEKYPHYQGGFIWDWID
                     QGLEKDGHLLYGGDFDDRPTDYEFCGNGLVFADRTESPKLANVKALYANLKLEVKDGQ
                     LFLKNDNLFTNSSSYYFLTSLLVDGKLTYQSRPLTFGLEPGESGTFALPWPEVADEKG
                     EVVYRVTAHLKEDLPWADEGFTVAEAEEVAQKLPEFKPEGRPDLVDSDYNLGLKGNNF
                     QILFSKVKGWPVSLKYAGREYLKRLPEFTFWRALTDNDRGAGYGYDLARWENAGKYAR
                     LKDISCEVKEDSVLVKTAFTLPVALKGDLTVTYEVDGRGKIAVTADFPGAEEAGLLPA
                     FGLNLALPKELTDYRYYGLGPNESYPDRLEGNYLGIYQGAVKKNFSPYLRPQETGNRS
                     KVRWYQLFDEKGGLEFTANGADLNLSALPYSAAQIEAADHAFELTNNYTWVRALSAQM
                     GVGGDDSWGQKVHPEFCLDAQKARQLRLVIQPLLLK"
BASE COUNT     1231 a   1322 c   1292 g   1170 t
ORIGIN
```

Query Match 41.7%; Score 3386.6; DB 3; Length 5015;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3408; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

```
Qy   4584 ccgataaaaagttaggcgacggctttgccctggtgccagcagacggtaaggtctacgcgc 4643
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1591 CCGACAAAAAGTTAGGCGACGGCTTTGCCCTGGTGCCAGCAGACGGTAAGGTCTACGCGC 1650

Qy   4644 catttgccggtactgtccgccagctggccaagacccggcactcgatcgtcctggaaaatg 4703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1651 CATTTGCCGGTACTGTCCGCCAGCTGGCCAAGACCCGGCACTCGATCGTCCTGGAAAATG 1710

Qy   4704 aacatggggtcttggtcttgattcaccttggcctgggcacggtcaaattaaacgggactg 4763
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   1711 AACATGGGGTCTTGGTCTTGATTCACCTTGGCCTGGGCACGGCCAAATTAAACGGGACTG 1770

Qy   4764 gctttgtcagctatgttgaagagggcagccaggtagaagccggccagcagatcctggaat 4823
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1771 GCTTTGTCAGCTATGTTGAAGAGGGCAGCCAGGTAGAAGCCGGCCAGCAGATCCTGGAAT 1830

Qy   4824 tctgggacccggcgatcaagcaggccaagctggacgacacggtaatcgtgaccgtcatca 4883
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1831 TCTGGGACCCGGCGATCAAGCAGGCCAAGCTGGACGACACGGTAATCGTGACCGTCATCA 1890

Qy   4884 acagcgaaactttcacaaatagccagatgctcttgccgatcggccacagcgtccaagccc 4943
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1891 ACAGCGAAACTTTCGCAAATAGCCAGATGCTCTTGCCGATCGGCCACAGCGTCCAAGCCC 1950

Qy   4944 tggatgatgtattcaagttagaagggaagaattagaaaatgagcaataagttagtaaaag 5003
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1951 TGGATGATGTATTCAAGTTAGAAGGGAAGAATTAGAAAATGAGCAATAAGTTAGTAAAAG 2010

Qy   5004 aaaaaagagttgaccaggcagacctggcctggctgactgacccggaagtttacgaagtca 5063
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2011 AAAAAAGAGTTGACCAGGCAGACCTGGCCTGGCTGACTGACCCGGAAGTTTACGAAGTCA 2070

Qy   5064 atacaattccccccgcactccgaccatgagtccttccaaagccaggaagaactggaggagg 5123
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2071 ATACAATTCCCCCCGCACTCCGACCATGAGTCCTTCCAAAGCCAGGAAGAACTGGAGGAGG 2130

Qy   5124 gcaagtccagtttagtgcagtccctggacggggactggctgattgactacgctgaaaacg 5183
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   2131 GCAAGTCCAGTTTAGTGCAGTCCTTGGACGGGGACTGGCTGATTGACTACGCTGAAAACG 2190

Qy   5184 gccagggaccagtcaacttctatgcagaagactttgacgatagcaattttaagtcagtca 5243
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2191 GCCAGGGACCAGTCAACTTCTATGCAGAAGACTTTGACGATAGCAATTTTAAGTCAGTCA 2250

Qy   5244 aagtacccggcaacctggaactgcaaggctttggccagccccagtatgtcaacgtccaat 5303
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2251 AAGTACCCGGCAACCTGGAACTGCAAGGCTTTGGCCAGCCCCAGTATGTCAACGTCCAAT 2310

Qy   5304 atccatgggacggcagtgaggagattttcccgccccaaattccaagcaaaaatccgctcg 5363
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2311 ATCCATGGGACGGCAGTGAGGAGATTTTCCCGCCCCAAATTCCAAGCAAAAATCCGCTCG 2370

Qy   5364 cttcttatgtcagatactttgacctggatgaagctttctgggacaaggaagtcagcttga 5423
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2371 CTTCTTATGTCAGATACTTTGACCTGGATGAAGCTTTCTGGGACAAGGAAGTCAGCTTGA 2430

Qy   5424 agtttgacggggcggcaacagccatctatgtctggctgaacggccacttcgtcggctacg 5483
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2431 AGTTTGACGGGGCGGCAACAGCCATCTATGTCTGGCTGAACGGCCACTTCGTCGGCTACG 2490

Qy   5484 gggaagactcctttaccccaagcgagtttatggttaccaagttcctcaagaaagaaaata 5543
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   2491 GGGAAGACTCCTTTACCCCAAGCGAGTTTATGGTTACCAAGTTCCTCAAGAAAGAAAACA 2550

Qy   5544 accgcctggcagtggctctctacaagtattcttccgcctcctggctggaagaccaggact 5603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2551 ACCGCCTGGCAGTGGCTCTCTACAAGTATTCTTCCGCCTCCTGGCTGGAAGACCAGGACT 2610

Qy   5604 tctggcgcatgtctggtttgttcagatcagtgactcttcaggccaagccgcgtctgcact 5663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2611 TCTGGCGCATGTCTGGTTTGTTCAGATCAGTGACTCTTCAGGCCAAGCCGCGTCTGCACT 2670

Qy   5664 tggaggaccttaagcttacggccagcttgaccgataactaccaaaaaggaaagctggaag 5723
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2671 TGGAGGACCTTAAGCTTACGGCCAGCTTGACCGATAACTACCAAAAAGGAAAGCTGGAAG 2730

Qy   5724 tcgaagccaatattgcctaccgcttgccaaatgccagctttaagctggaagtgcgggata 5783
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2731 TCGAAGCCAATATTGCCTACCGCTTGCCAAATGCCAGCTTTAAGCTGGAAGTGCGGGATA 2790

Qy   5784 gtgaaggtgacttggttgctgaaaagctgggcccaatcagaagcgagcagctggaattca 5843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2791 GTGAAGGTGACTTGGTTGCTGAAAAGCTGGGCCCAATCAGAAGCGAGCAGCTGGAATTCA 2850

Qy   5844 ctctggctgatttgccagtagctgcctggagcgcggaaaagcctaacctttaccaggtcc 5903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2851 CTCTGGCTGATTTGCCAGTAGCTGCCTGGAGCGCGGAAAAGCCTAACCTTTACCAGGTCC 2910

Qy   5904 gcctgtatttataccaggcaggcagcctcttagaggttagccggcaggaagtgggtttcc 5963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2911 GCCTGTATTTATACCAGGCAGGCAGCCTCTTAGAGGTTAGCCGGCAGGAAGTGGGTTTCC 2970

Qy   5964 gcaactttgaactaaaagacgggattatgtaccttaacggccagcggatcgtcttcaagg 6023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2971 GCAACTTTGAACTAAAAGACGGGATTATGTACCTTAACGGCCAGCGGATCGTCTTCAAGG 3030

Qy   6024 gggccaaccggcacgaatttgacagtaagttgggtcgggctatcacggaagaggatatga 6083
          |||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||
Db   3031 GGGCCAACCGGCACGAATTTGACAGTAAGTTGGGCCGGGCTATCACAGAAGAGGATATGA 3090

Qy   6084 tctgggacatcaagaccatgaagcgaagcaacatcaatgctgtccgctgctctctcactacc 6143
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3091 TCTGGGATATCAAGACCATGAAGCGAAGCAACATCAATGCTGTCCGCTGCTCTCACTACC 3150

Qy   6144 cgaaccagtccctcttttaccggctctgtgacaagtacggcctttacgtcattgatgaag 6203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3151 CGAACCAGTCCCTCTTTTACCGGCTCTGTGACAAGTACGGCCTTTACGTCATTGATGAAG 3210

Qy   6204 ctaacctggaaagccacggcacctgggaaaaagtgggggggcacgaagatcctagcttca 6263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3211 CTAACCTGGAAAGCCACGGCACCTGGGAAAAAGTGGGGGGGCACGAAGATCCTAGCTTCA 3270

Qy   6264 atgttccaggcgatgaccagcattggctgggagccagcttatcccgggtgaagaacatga 6323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3271 ATGTTCCAGGCGATGACCAGCATTGGCTGGGAGCCAGCTTATCCCGGGTGAAGAACATGA 3330

Qy   6324 tggctcgggacaagaaccatgcttcaatcctaatctggtctttaggcaatgagtcttacg 6383
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3331 TGGCTCGGGACAAGAACCATGCTTCAATCCTGATCTGGTCTTTAGGCAATGAGTCTTACG 3390

Qy   6384 ccggcactgtctttgcccaaatggctgattacgtccggaaggctgatccgacccgggttc 6443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3391 CCGGCACTGTCTTTGCCCAAATGGCTGATTACGTCCGGAAGGCTGATCCGACCCGGGTTC 3450

Qy   6444 agcactatgaaggggtgacccacaaccggaagtttgacgacgccacccagattgaaagcc 6503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3451 AGCACTATGAAGGGGTGACCCACAACCGGAAGTTTGACGACGCCACCCAGATTGAAAGCC 3510

Qy   6504 ggatgtatgctccggccaaggtaattgaagaatacttgaccaataaaccagccaagccat 6563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3511 GGATGTATGCTCCGGCCAAGGTAATTGAAGAATACTTGACCAATAAACCAGCCAAGCCAT 3570

Qy   6564 ttatctcagttgaatacgctcacgccatgggcaactccgtcggtgacctggccgcctaca 6623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3571 TTATCTCAGTTGAATACGCTCACGCCATGGGCAACTCCGTCGGTGACCTGGCCGCCTACA 3630

Qy   6624 cggccctggaaaaataccccccactaccagggcggcttcatctgggactggattgaccaag 6683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3631 CGGCCCTGGAAAAATACCCCCCACTACCAGGGCGGCTTCATCTGGGACTGGATTGACCAAG 3690

Qy   6684 gactggaaaaagacgggcacctgctttatgggggcgacttcgatgaccggccaaccgact 6743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3691 GACTGGAAAAAGACGGGCACCTGCTTTATGGGGGCGACTTCGATGACCGGCCAACCGACT 3750

Qy   6744 atgaattctgcgggaacggcctggtctttgctgaccggactgaatcgccgaaactggcta 6803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3751 ATGAATTCTGCGGGAACGGCCTGGTCTTTGCTGACCGGACTGAATCGCCGAAACTGGCTA 3810

Qy   6804 atgtcaaggccctttacgccaaccttaagttagaagtaaaagatgggcagctcttcctca 6863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3811 ATGTCAAGGCCCTTTACGCCAACCTTAAGTTAGAAGTAAAAGATGGGCAGCTCTTCCTCA 3870

Qy   6864 aaaacgacaatttatttaccaacagctcatcttactacttcttgactagtcttttggtcg 6923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3871 AAAACGACAATTTATTTACCAACAGCTCATCTTACTACTTCTTGACTAGTCTTTTGGTCG 3930

Qy   6924 atggcaagttgacctaccagagccggcctctgacctttggcctggagcctggcgaatccg 6983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3931 ATGGCAAGTTGACCTACCAGAGCCGGCCTCTGACCTTTGGCCTGGAGCCTGGCGAATCCG 3990

Qy   6984 ggacctttgccctgccttggccggaagtcgctgatgaaaaaggggaggtcgtctaccggg 7043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3991 GGACCTTTGCCCTGCCTTGGCCGGAAGTCGCTGATGAAAAAGGAGAGGTCGTCTACCGGG 4050

Qy   7044 taacggcccacttaaaagaagacttgccttgggcggatgagggcttcactgtggctgaag 7103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4051 TAACGGCCCACTTAAAAGAAGACTTGCCTTGGGCGGATGAGGGCTTCACTGTGGCTGAAG 4110

Qy   7104 cagaagaagtagctcaaaagctgccggaatttaagccggaagggcggccagatttagttg 7163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4111 CAGAAGAAGTAGCTCAAAAGCTGCCGGAATTTAAGCCGGAAGGGCGGCCAGATTTAGTTG 4170

Qy   7164 attccgactacaacctaggcctgaaaggaaataacttccaaattctcttctccaaggtca 7223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4171 ATTCCGACTACAACCTAGGCCTGAAAGGAAATAACTTCCAAATTCTCTTCTCCAAGGTCA 4230

Qy   7224 agggctggccggtttccctcaagtatgccggtagggaatacttgaagcggctgccggaat 7283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4231 AGGGCTGGCCGGTTTCCCTCAAGTATGCCGGTAGGGAATACTTGAAGCGGCTGCCGGAAT 4290

Qy   7284 ttaccttctggcgggccctgacggacaacgaccggggagctggttacggctatgatctgg 7343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4291 TTACCTTCTGGCGGGCCCTGACGGACAACGACCGGGGAGCTGGTTACGGCTATGATCTGG 4350

Qy   7344 cccggtgggaaaatgccggcaagtatgcccgcttgaaagacatcagctgcgaggtcaagg 7403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4351 CCCGGTGGGAAAATGCCGGCAAGTATGCCCGCTTGAAAGACATCAGCTGCGAGGTCAAGG 4410

Qy   7404 aagactccgttttggtcaagactgcctttacgttgcctgtcgccttaaagggtgatttaa 7463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4411 AAGACTCCGTTTTGGTCAAGACTGCCTTTACGTTGCCTGTCGCCTTAAAGGGTGATTTAA 4470

Qy   7464 ccgtgacctatgaagtcgatggacggggcaagattgctgtaacagctgacttcccaggcg 7523
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4471 CCGTGACCTATGAAGTCGATGGACGGGGCAAGATTGCTGTAACAGCTGACTTCCCAGGCG 4530

Qy   7524 cggaagaagctggtctcttgccagcctttggcttgaacctggccctgccaaaagaactga 7583
          |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||||
Db   4531 CGGAAGAAGCCGGTCTCTTGCCAGCCTTTGGCTTGAACCTAGCCCTGCCAAAAGAACTGA 4590

Qy   7584 ccgattaccgctactatggtctgggacctaatgagagctacccagaccgcttggaaggta 7643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4591 CCGATTACCGCTACTATGGTCTGGGACCTAATGAGAGCTACCCAGACCGCTTGGAAGGTA 4650

Qy   7644 attacctgggcatctaccagggagcggtaaaaaagaacttagcccata---tcgtccgc 7700
          |||||||||||||||||||||||||||||||||||||||||||||||    |||||||
Db   4651 ATTACCTGGGCATCTACCAGGGAGCGGTAAAAAAGAACTTTAGCCCATACCTGCGTCCGC 4710

Qy   7701 aggaaacgggcaaccggagcaaggttcgctggtaccagctctttgatgaaaagggcggct 7760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4711 AGGAAACGGGCAACCGGAGCAAGGTTCGCTGGTACCAGCTCTTTGATGAAAAGGGCGGCT 4770

Qy   7761 tggaatttacggccaatggggcagacttgaacttgtctgctttgccatattctgccgccc 7820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4771 TGGAATTTACGGCCAATGGGGCAGACTTGAACTTGTCTGCTTTGCCATATTCTGCCGCCC 4830

Qy   7821 aaattgaagcagcggaccacgcttttgaactgactaacaattacacttgggttagagcct 7880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4831 AAATTGAAGCAGCGGACCACGCTTTTGAACTGACTAACAATTACACTTGGGTTAGAGCCT 4890

Qy   7881 taagcgcccagatgggggtcggcggggatgactcctgggggcagaaggtccacccggaat 7940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4891 TAAGCGCCCAGATGGGGGTCGGCGGGGATGACTCCTGGGGGCAGAAGGTCCACCCGGAAT 4950

Qy   7941 tctgcctggatgctcaaaaagcccgccagcttcgcctggtgattcagccccttttactaa 8000
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   4951 TCTGCCTGGATGCTCAAAAAGCCCGCCAGCTCCGCCTGGTGATTCAGCCCCTTTTACTAA 5010

Qy   8001 aataa 8005
          |||||
Db   5011 AATAA 5015
```

```
RESULT  4
EVE132038
LOCUS       EVE132038    3840 bp    mRNA  circular  SYN       28-JUL-1999
DEFINITION  Expression vector pCDPT.
ACCESSION   AJ132038
VERSION     AJ132038.1  GI:5640088
KEYWORDS    AMP gene; beta lactamase; ColE1 origin of replication; multiple
            cloning site; SP6 promoter; SV40 origin of replication; T7
            promoter; xanthine-guanine phosphoribosyl transferase;
            Xanthine-guanine phosphoribosyl transferase gene.
SOURCE      Expression vector pCDPT.
  ORGANISM  Expression vector pCDPT
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 3840)
  AUTHORS   Zeng,B.J.
  TITLE     Mammalian Expression Vector for with fuse Xanthine-guanine
            phosphoribosyl transferase Tag
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3840)
  AUTHORS   Zeng,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center,
            Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080,
            CHINA
FEATURES             Location/Qualifiers
     source          1. .3840
                     /organism="Expression vector pCDPT"
                     /db_xref="taxon:90749"
     promoter        209. .863
                     /note="CMV"
     promoter        864. .882
                     /note="T7"
     misc_feature    882. .984
                     /note="Multiple cloning site; HindIII, BamHI, BstXI,
                     EcoRI, NotI, XhoI"
     CDS             929. .1387
                     /codon_start=1
                     /product="Xanthine-guanine phosphoribosyl transferase"
                     /protein_id="CAB51567.1"
                     /db_xref="GI:5640089"
                     /translation="MSEKYIVTWDMLQIHARKLASRLMPSEQWKGIIAVSRGGLVPGA
                     LLARELGIRHVDTVCISSYDHDNQRELKVLKRAEGDGEGFIVIDDLVDTGGTAVAIRE
                     MYPKAHFVTIFAKPAGRPLVDDYVVDIPQDTWIEQPWDMGVVFVPPISGR"
     polyA_site      1649. .1863
                     /note="BGH"
     promoter        2450. .2775
                     /note="SP6"
     rep_origin      2644. .2729
                     /note="SV40"
     gene            complement(2844. .3704)
                     /gene="amp"
     CDS             complement(2844. .3704)
                     /gene="amp"
                     /codon_start=1
                     /product="beta-lactamase"
```

```
                     /protein_id="CAB51568.1"
                     /db_xref="GI:5640090"
                     /translation="MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
                     IELDLNSGKILESFRPEERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
                     YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTTIGGPKELTAFLHNMGDHVTRL
                     DRWEPELNEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAGPL
                     LRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIA
                     EIGASLIKHW"
     rep_origin      3632. .>3840
                     /note="ColE1"
BASE COUNT      938 a    960 c    964 g    978 t
ORIGIN


  Query Match           16.5%;  Score 1336.4;  DB 56;  Length 3840;
  Best Local Similarity  75.4%;  Pred. No. 9.2e-309;
  Matches 1991;  Conservative    0;  Mismatches   16;  Indels  635;  Gaps    1;

Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 GATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCA 194

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA 254

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT 314

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGG 374

Qy    241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC 434

Qy    301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT 494

Qy    361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG 554

Qy    421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    555 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC 614

Qy    481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    615 ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT 674

Qy    541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    675 AACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA 734

Qy    601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    735 AGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACG 794

Qy    661 actcactatagggagacccaagcttggtaccgagctcggatccactagtaacggccgcca 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    795 ACTCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCA 854

Qy    721 gtgtgctggaattctgcagatatccatcacactggcggccgctcgagcatgcatctaga- 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    855 GTGTGCTGGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAG 914

Qy    780 ------------------------------------------------------------ 779

Db    915 CTGGGACACTTCACATGAGCGAAAAATACATCGTCACCTGGGACATGTTGCAGATCCATG 974

Qy    780 ------------------------------------------------------------ 779

Db    975 CACGTAAACTCGCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAA 1034

Qy    780 ------------------------------------------------------------ 779

Db   1035 GCCGTGGCGGTCTGGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCG 1094

Qy    780 ------------------------------------------------------------ 779

Db   1095 ATACCGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAAAC 1154

Qy    780 ------------------------------------------------------------ 779

Db   1155 GCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTA 1214

Qy    780 ------------------------------------------------------------ 779

Db   1215 CTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCAAAAC 1274

Qy    780 ------------------------------------------------------------ 779

Db   1275 CGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAGATACCTGGATTG 1334

Qy    780 ------------------------------------------------------------ 779

Db   1335 AACAGCCGTGGGATATGGGCGTCGTATTCGTCCCGCCAATCTCCGGTCGCTAATCTTTTC 1394

Qy    780 ------------------------------------------------------------ 779

Db   1395 AACGCCTGGCACTGCCGGGCGTTGTTCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCA 1454

Qy    780 ------------------------------------------------------------ 779

Db   1455 GTAAGTATTCTGGAGGCTGCATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAA 1514

Qy    780 ----------------------------------gggccctattctatagtgtcacctaa 805
                                               ||||||||||||||||||||||||||
Db   1515 CCCCGCTTTAAACATCCTGAAACCTCGACGCTAGGGGCCCTATTCTATAGTGTCACCTAA 1574

Qy    806 atgctagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgttt 865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1575 ATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTT 1634

Qy    866 gcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaat 925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1635 GCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT 1694

Qy    926 aaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtgggg 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1695 AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGG 1754

Qy    986 tggggcaggacagcaaggggggaggattgggaagacaatagcaggcatgctggggatgcgg 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1755 TGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGG 1814

Qy   1046 tgggctctatggcttctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgc 1105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1815 TGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGCATTAATGAATCGGCCAACGCGC 1874

Qy   1106 ggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcg 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1875 GGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG 1934

Qy   1166 ctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatc 1225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1935 CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATC 1994

Qy   1226 cacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccag 1285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1995 CACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG 2054
```

```
Qy   1286 gaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccctgacgagca 1345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2055 GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA 2114

Qy   1346 tcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagatacca 1405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2115 TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCA 2174

Qy   1406 ggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccgg 1465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2175 GGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGG 2234

Qy   1466 atacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtag 1525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2235 ATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAG 2294

Qy   1526 gtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgt 1585
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2295 GTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT 2354

Qy   1586 tcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagaca 1645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2355 TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA 2414

Qy   1646 cgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtagg 1705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2415 CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG 2474

Qy   1706 cggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatt 1765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2475 CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT 2534

Qy   1766 tggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatc 1825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2535 TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATC 2594

Qy   1826 cggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcagattacgcg 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2595 CGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCG 2654

Qy   1886 cagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtg 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2655 CAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG 2714

Qy   1946 gaacgaaaactcacgttaagggattttggtcatgagcggatacatatttgaatgtattta 2005
          ||||||||||||||||||||||||||||||||||  ||  | |     |  |  ||
Db   2715 GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTA 2774

Qy   2006 ga 2007
          ||
Db   2775 GA 2776
```

```
RESULT   5
AR098190
LOCUS       AR098190     3853 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 5 from patent US 6074850.
ACCESSION   AR098190
VERSION     AR098190.1  GI:12807447
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 3853)
  AUTHORS   Antelman,D., Gregory,R.J. and Wills,K.N.
  TITLE     Retinoblastoma fusion polypeptides
  JOURNAL   Patent: US 6074850-A 5 13-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .3853
                     /organism="unknown"
BASE COUNT      936 a    987 c    941 g    989 t
ORIGIN
```

```
 Query Match             15.3%;  Score 1240.4;  DB 9;  Length 3853;
 Best Local Similarity   74.1%;  Pred. No. 8.8e-286;
 Matches 1911;  Conservative    0;  Mismatches   96;  Indels  573;  Gaps    4;

Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 GATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCA 269

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA 329

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT 389

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGG 449

Qy    241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC 509

Qy    301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT 569

Qy    361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG 629

Qy    421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC 689

Qy    481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT 749

Qy    541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
          ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db    750 AACAACTCCGCCCCATTGACGCAAATGGGCGGTAGCGCTGTACGGTGGGAGGTCTATATA 809

Qy    601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 AGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACG 869

Qy    661 actcactatagggagacccaagctt---------------------------------- 685
          |||||||||||||||||||||||||
Db    870 ACTCACTATAGGGAGACCCAAGCTTCGCGCGGGTACCACTCTCTTCCGCATCGCTGTCTG 929

Qy    686 ---------------------------------------------------------- 685

Db    930 CGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTC 989

Qy    686 -------------------ggtaccgagctcggatccactagtaacggccgccagtgtgc 726
                              |   |||| |     ||| | |   | || |   || | |
Db    990 TTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGT 1049

Qy    727 tggaattctgcagatatccatcacactggcggccgc-------------------tcgag 767
            | ||    | |||    |   | || | ||| ||                   ||| |
Db   1050 CCGCATCGACCGGATCGGAAAACCTCTCGAGGCGGCCGCTGCAGTCTAGACGAATTCGCG 1109

Qy    768 catgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctgatcagc 827
            | |        ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 TACGATATCGATGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGC 1169

Qy    828 ctcgactgtgccttctagttgccagccatctgttgtttgcccctccccccgtgccttcctt 887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 CTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTT 1229
```

```
Qy    888 gaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgca 947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 GACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCA 1289

Qy    948 ttgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaaggggga 1007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 TTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA 1349

Qy   1008 ggattgggaagacaatagcaggcatgctgggga---------------------------- 1040
          ||||||||||||||||||  |  |||   |
Db   1350 GGATTGGGAAGACAATAGCCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAG 1409

Qy   1041 ------------------------------------------------------------ 1040

Db   1410 ATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACG 1469

Qy   1041 ------------------------------------------------------------ 1040

Db   1470 CCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACT 1529

Qy   1041 ------------------------------------------------------------ 1040

Db   1530 TGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATA 1589

Qy   1041 ------------------------------------------------------------ 1040

Db   1590 AAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATC 1649

Qy   1041 ------------------------------------------------------------ 1040

Db   1650 ATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTC 1709

Qy   1041 ------------------------------------------------------------ 1040

Db   1710 CTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGT 1769

Qy   1041 ----------------------------------------------tgcggtgggctctatgg 1057
                                                        ||||  ||  ||||  ||
Db   1770 GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC 1829

Qy   1058 cttct----------gaggcggaaagaaccagctgcattaatgaatcggccaacgcgcgg 1107
          |  ||          ||  |     |  ||||||||||||||||||||||||||||||
Db   1830 CCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGG 1889

Qy   1108 ggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgct 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1890 GGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT 1949

Qy   1168 cggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatcca 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1950 CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA 2009

Qy   1228 cagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccagga 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2010 CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGA 2069

Qy   1288 accgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatc 1347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2070 ACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC 2129

Qy   1348 acaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2130 ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG 2189

Qy   1408 cgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggat 1467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2190 CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGAT 2249

Qy   1468 acctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggt 1527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2250 ACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGT 2309

Qy   1528 atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttc 1587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2310 ATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTC 2369

Qy   1588 agcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacg 1647
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2370 AGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG 2429

Qy   1648 acttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg 1707
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2430 ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG 2489

Qy   1708 gtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttg 1767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2490 GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTG 2549

Qy   1768 gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccg 1827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2550 GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG 2609

Qy   1828 gcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcagcagattacgcgca 1887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2610 GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA 2669

Qy   1888 gaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtgga 1947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2670 GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA 2729

Qy   1948 acgaaaactcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          ||||||||||||||||||||||||||||||||||  ||  | |       |  |  ||||
Db   2730 ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 2789
```

```
RESULT   6
LP1PLPREP/c
LOCUS       LP1PLPREP    2093 bp    DNA             BCT       24-APR-1996
DEFINITION  Plasmid pLP1 (from Lactobacillus plantarum CCM 1904) replication
            protein (rep) gene, complete cds.
ACCESSION   M31223
VERSION     M31223.1  GI:149684
KEYWORDS    replication protein.
SOURCE      Plasmid pLP1 (tissue library: CCM 1904) DNA.
  ORGANISM  Plasmid Plasmid pLP1
            plasmids.
REFERENCE   1  (bases 1 to 2093)
  AUTHORS   Bouia,A., Bringel,F., Frey,L., Kammerer,B., Belarbi,A.,
            Guyonvarch,A. and Hubert,J.C.
  TITLE     Structural organization of pLP1, a cryptic plasmid from
            Lactobacillus plantarum CCM 1904
  JOURNAL   Plasmid 22 (3), 185-192 (1989)
  MEDLINE   90222315
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by B.Kammerer, 11-JAN-1990.
FEATURES             Location/Qualifiers
     source          1. .2093
                     /organism="Plasmid pLP1"
                     /plasmid="Plasmid pLP1"
                     /specific_host="Lactobacillus plantarum"
                     /db_xref="taxon:2593"
                     /tissue_lib="CCM 1904"
     gene            249. .1215
                     /gene="rep"
     RBS             249. .253
                     /gene="rep"
     CDS             262. .1215
                     /gene="rep"
                     /codon_start=1
                     /transl_table=11
                     /product="replication protein"
                     /protein_id="AAA98164.1"
                     /db_xref="GI:149685"
                     /translation="MSEIFEDKTENGKVRPWRERKIENVRYAEYLAILEFKRAHDVRG
                     CGEVLRFRKIGEHLKLYQTWFCHKRLCPLCNWRRSMKNSSQLKQIIAEAVAREPKGRF
                     LFLTLTVKNAHSAEELKVSLRALTKAFNKLTRYKKVTKNLLGYLRSTEITVNEQDGSY
```

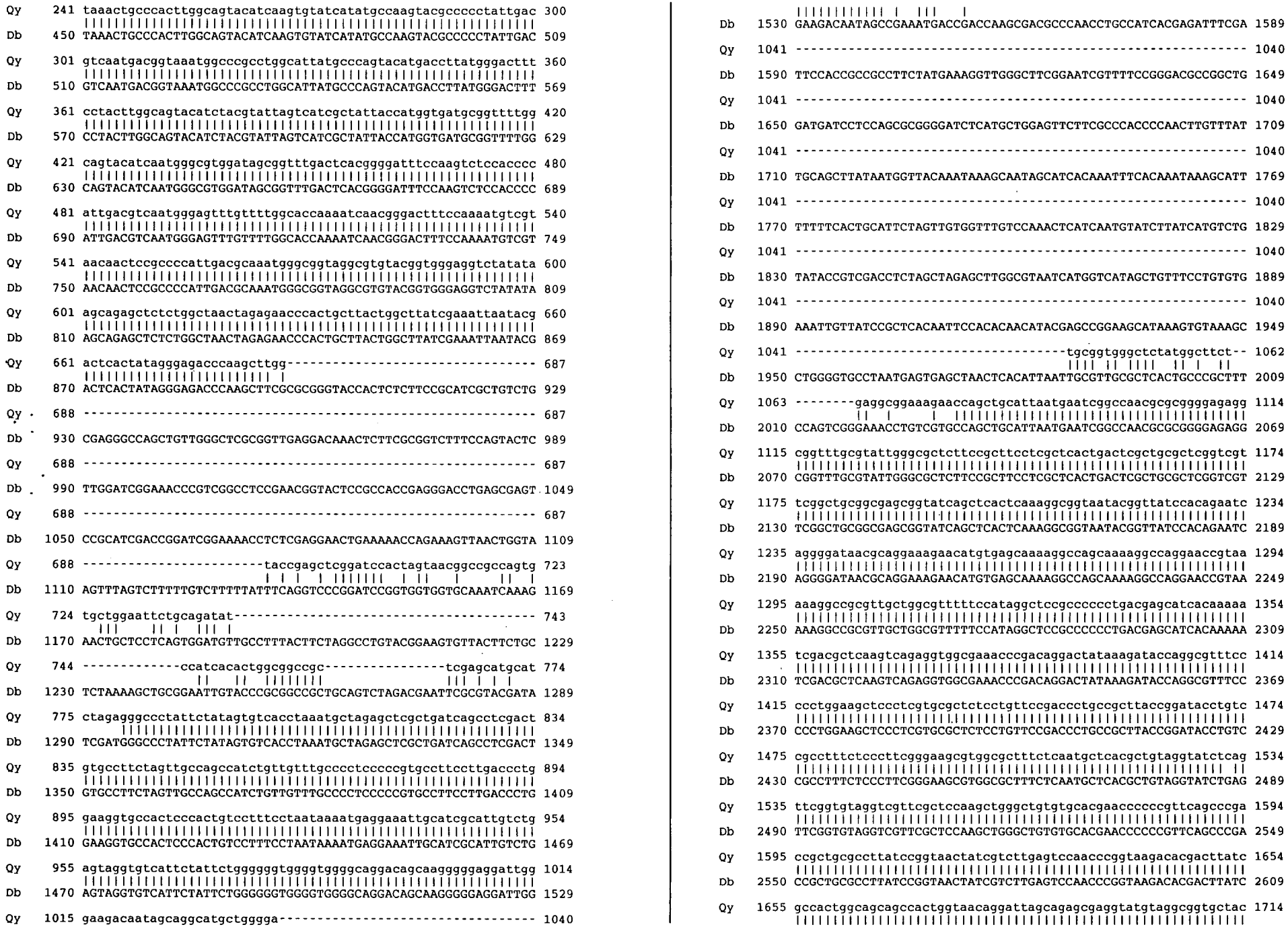
```
                     NQHLHVLLFVKSSYFKNSNNYLAQAEWAKLWQKALKVDYEPVVHVQAVKANKRKGTDS
                     LQASAEETAKYEVKSADYMTADDERNLVVIKNLEYALAGTRQISYGGLLKQIKQDLKL
                     EDVENGDLVHVGDEDYTKEQMEAAEEVVAKWDFNKQNYFIW"
     repeat_region   1335. .1351
                     /note="direct repeat A"
     repeat_region   1352. .1368
                     /note="direct repeat B"
     repeat_region   1369. .1385
                     /note="direct repeat C"
     repeat_region   1386. .1402
                     /note="direct repeat D"
     repeat_region   1403. .1419
                     /note="direct repeat E"
     repeat_region   1420. .1436
                     /note="direct repeat F"
     repeat_region   1437. .1453
                     /note="direct repeat G"
     repeat_region   1456. .1467
                     /note="direct repeat H"
     repeat_region   1468. .1484
                     /note="direct repeat I"
     repeat_region   1485. .1501
                     /note="direct repeat J"
     repeat_region   1504. .1520
                     /note="direct repeat K"
     repeat_region   1521. .1537
                     /note="direct repeat L"
     repeat_region   1538. .1554
                     /note="direct repeat M"
     rep_origin      2051. .2064
                     /note="origin of replication (put.); putative"
BASE COUNT      639 a    335 c    466 g    653 t
ORIGIN      1 bp upstream of HindIII site.


 Query Match             15.3%;  Score 1238.4;  DB 3;  Length 2093;
 Best Local Similarity   98.0%;  Pred. No. 2.6e-285;
 Matches 1254;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy   2032 gcacatttccccgaaaagtgccacctgacgtcgacggatcgggagatcatatcctgacat 2091
          |||   |  |   ||||  |  |  ||  |  ||       | |  ||||||||||||||||
Db   1280 GCAAGCTGACAACAAAAAAGGGATATGCCAACGGATTTACCGTTGATCATATCCTGACAT 1221

Qy   2092 tctctttaccaaataaaataatttttgtttattaaaatcccatttgcgacaacttcttcc 2151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1220 TCTCTTTACCAAATAAAATAATTTTGTTTATTAAAATCCCATTTTGCGACAACTTCTTCC 1161

Qy   2152 gcagcttccatttgctctttggtgtaatcttcatcgccaacatgaactaaatcaccattc 2211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1160 GCAGCTTCCATTTGCTCTTTGGTGTAATCTTCATCGCCAACATGAACTAAATCACCATTC 1101

Qy   2212 tcaacatcttcaagtttcaaatcttgcttaatttgctttaataatccaccatagctgatt 2271
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1100 TCAACATCTTCAAGTTTCAAATCTTGCTTAATTTGCTTTAATAATCCACCATAGCTGATT 1041

Qy   2272 tgtcgtgttccagctaaggcatactccaaatttttaatcaccaccaaattacgctcatca 2331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1040 TGTCGTGTTCCAGCTAAGGCATACTCCAAATTTTTAATCACCACCAAATTACGCTCATCA 981

Qy   2332 tcagccgtcatataatcagctgattttacctcgtatttcgccgtttcttcggcactagct 2391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    980 TCAGCCGTCATATAATCAGCTGATTTTACCTCGTATTTCGCCGTTTCTTCGGCACTAGCT 921

Qy   2392 tgcaaagagtcagttcctttacgtttgttagctttaacagcctgcacatgcaccacaggc 2451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    920 TGCAAAGAGTCAGTTCCTTTACGTTTGTTAGCTTTAACAGCCTGCACATGCACCACAGGC 861

Qy   2452 tcataatcaactttcaaggctttttgccataatttttgcccattctgcttgtgctaaataa 2511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    860 TCATAATCAACTTTCAAGGCTTTTTGCCATAATTTTGCCCATTCTGCTTGTGCTAAATAA 801

Qy   2512 ttatttgaattcttaaaataacttgattttacaaacagcaacacatgcaagtgttgatta 2571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    800 TTATTTGAATTCTTAAAATAACTTGATTTTACAAACAGCAACACATGCAAGTGTTGATTA 741

Qy   2572 tatgacccgtcttgttcattaacggtaatttccgttgaacgtaaataacccaataaattt 2631
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    740 TATGACCCGTCTTGTTCATTAACGGTAATTTCCGTTGAACGTAAATAACCCAATAAATTT 681

Qy   2632 ttagtcacttttttatagcgagttagcttattaaaggctttagtcaaagctcttaaagac 2691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    680 TTAGTCACTTTTTTATAGCGAGTTAGCTTATTAAAGGCTTTAGTCAAAGCTCTTAAAGAC 621

Qy   2692 acttttaactcctctgctgaatgagcgtttttaacggttaaagttaaaaacaaaaaccgt 2751
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    620 ACTTTTAACTCCTCTGCTGAATGAGCGTTTTTAACGGTTAAAGTTAAAAACAAAAACCGT 561

Qy   2752 cctttaggctctcttgcaactgcttccgcaataatttgttttaactggctcgagtttttc 2811
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560 CCTTTAGGCTCTCTTGCAACTGCTTCCGCAATAATTTGTTTTAACTGGCTCGAGTTTTTC 501

Qy   2812 atgctccttctccaattacacaatggacacaatcgtttatgacaaaaccacgtttgataa 2871
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    500 ATGCTCCTTCTCCAATTACACAATGGACACAATCGTTTATGACAAAACCACGTTTGATAA 441

Qy   2872 agttttaagtgctcgccaatcttacgaaaacgcaaaacttcaccacaaccccgtacatca 2931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    440 AGTTTTAAGTGCTCGCCAATCTTACGAAAACGCAAAACTTCACCACAACCCCGTACATCA 381

Qy   2932 tgtgcccgtttaaattctaagattgccaaatattcggcatagcgcacattttcaatcttc 2991
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    380 TGTGCCCGTTTAAATTCTAAGATTGCCAAATATTCGGCATAGCGCACATTTTCAATCTTC 321

Qy   2992 cgttctcgccaaggtctaactttgccattttcagttttatcttcaaaaatttctgacata 3051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    320 CGTTCTCGCCAAGGTCTAACTTTGCCATTTTCAGTTTTATCTTCAAAAATTTCTGACATA 261

Qy   3052 aaaagctcctccagtttatccacgtgaaggagctgactatctttttcaataagcttataa 3111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    260 AAAAGCTCCTCCAGTTTATCCACGTGAAGGAGCTGACTATCTTTTTCAATAAGCTTATAA 201

Qy   3112 ccttgacatcatagggcttttcccctagaataggctataaatcgcaaatgataatcaact 3171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 CCTTGACATCATAGGGCTTTTCCCCTAGAATAGGCTATAAATCGCAAATGATAATCAACT 141

Qy   3172 cacgtgttccgagcggccaaactaggaatttgcacgtgggtttttatttttgtctttcttt 3231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 CACGTGTTCCGAGCGGCCAAACTAGGAATTTGCACGTGGGTTTTTATTTTGTCTTTCTTT 81

Qy   3232 caaccaatttataaccctaataatacaccaaaagcctataaaatcaatggatacaagccc 3291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 CAACCAATTTATAACCCTAATAATACACCAAAAGCCTATAAAATCAATGGATACAAGCCC 21

Qy   3292 aattaagcctaatcaagctt 3311
          |||||||||||||||||||
Db     20 AATTAAGCCTAATCAAGCTT 1


RESULT   7
C30ORFR/c
LOCUS       C30ORFR      2140 bp    DNA             BCT       07-MAR-1996
DEFINITION  Plasmid pC30il (from L.plantarum) ORF R, complete genome.
ACCESSION   J03319
VERSION     J03319.1  GI:144180
KEYWORDS    complete genome.
SOURCE      Plasmid pC30il DNA.
  ORGANISM  Plasmid Plasmid pC30il
            plasmids.
REFERENCE   1  (bases 1 to 2140)
  AUTHORS   Skaugen,M.
  TITLE     The complete nucleotide sequence of a small cryptic plasmid from
            Lactobacillus plantarum
  JOURNAL   Plasmid 22 (2), 175-179 (1989)
  MEDLINE   90160873
COMMENT     Draft entry and computer-readable sequence for [1] kindly submitted
            by M.Skaugen, 19-NOV-1989.
```

```
FEATURES             Location/Qualifiers
     source          1. .2140
                     /organism="Plasmid pC30il"
                     /plasmid="Plasmid pC30il"
                     /specific_host="Lactobacillus plantarum"
                     /db_xref="taxon:2535"
     rep_origin      63. .85
                     /note="plus-strand origin of replication"
     -35_signal      145. .150
     -10_signal      167. .172
     RBS             361. .366
                     /note="ribosome binding site"
     CDS             372. .1325
                     /note="ORF R"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA91566.1"
                     /db_xref="GI:144181"
                     /translation="MSEIFEDKTENGKVRPWRERKIENVRYAEYLAILEFKRAHDVRG
                     CGEVLRFRKIGEHLKLYQTWFCHKRLCPLCNWRKSMKNSSQLKQIIAEAVAREPKGRF
                     LFLTLTVKNAHSAEELKVSLRALTKAFNKLTRYKKVTKNLLGYLRSTEITVNEQDGSY
                     NQHLHVLVFVKSSYFKNSNNYLAQAEWAKLWQRALKVDYEPVVHVQAVKANKRKGTDS
                     LQASAEETAKYEVKSADYMTADDERNLVVIKNLEYALAGTRQISYGGLFGQIKQDLKL
                     EDVENGDLVHVGDEDYTKEQMEAAEEVVAKWDFNKQNYFIW"
     misc_signal     1337. .1377
                     /note="transcription termination signal"
     repeat_region   1469. .1706
                     /note="tandem repeat region"
BASE COUNT      640 a    332 c    459 g    709 t
ORIGIN
```

```
  Query Match             14.7%;  Score 1193.8;  DB 2;  Length 2140;
  Best Local Similarity   93.0%;  Pred. No. 1.2e-274;
  Matches 1295;  Conservative    0;  Mismatches   92;  Indels    6;  Gaps    4;

Qy   2032 gcacatttccccgaaaagtgccacctgacgtcgacggatcgggagatcatatcctgacat 2091
          ||||   |  |   ||||  |  |  || |  ||      | |  ||||||||||||||||
Db   1390 GCAAGCTGACAACAAAAAAGGGATATGCCAACGGATTTACCGTTGATCATATCCTGACAT 1331

Qy   2092 tctctttaccaaataaaataatttgtttattaaaatcccatttgcgacaacttcttcc 2151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1330 TCTCTTTACCAAATAAAATAATTTTGTTTATTAAAATCCCATTTGCGACAACTTCTTCC 1271

Qy   2152 gcagcttccatttgctctttggtgtaatcttcatcgccaacatgaactaaatcaccattc 2211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1270 GCAGCTTCCATTTGCTCTTTGGTGTAATCTTCATCGCCAACATGAACTAAATCACCATTT 1211

Qy   2212 tcaacatcttcaagtttcaaatcttgcttaatttgctttaataatccaccatagctgatt 2271
          ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db   1210 TCAACATCTTCAAGTTTCAAATCTTGCTTAATTTGCCCAAATAATCCACCATAGCTGATT 1151

Qy   2272 tgtcgtgttccagctaaggcatactccaaatttttaatcaccaccaaattacgctcatca 2331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1150 TGTCGTGTTCCAGCTAAGGCATACTCCAAATTTTTAATCACCACCAAATTACGCTCATCA 1091

Qy   2332 tcagccgtcatataatcagctgattttacctcgtatttcgccgtttcttcggcactagct 2391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1090 TCAGCCGTCATATAATCAGCTGATTTTACCTCGTATTTCGCCGTTTCTTCGGCACTAGCT 1031

Qy   2392 tgcaaagagtcagttcctttacgtttgttagctttaacagcctgcacatgcaccacaggc 2451
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1030 TGCAAAGAGTCAGTTCCTTTGCGTTTGTTAGCTTTAACAGCCTGCACATGCACCACAGGC 971

Qy   2452 tcataatcaacttttcaaggctttttgccataattttgcccattctgcttgtgctaaataa 2511
          ||||||||||||| || ||| ||||||||||||||||||||||||||||||||||||||||
Db    970 TCATAATCAACTTTTAAAGCTCTTTGCCATAATTTTGCCCATTCTGCTTGTGCTAAATAA 911

Qy   2512 ttatttgaattcttaaaataacttgattttacaaacagcaacacatgcaagtgttgatta 2571
          ||||||||||||||||||||||   ||| ||||| | ||| ||||| || || ||||||
Db    910 TTATTTGAATTCTTAAAATAACTACTTTTCACAAAAACCAAAACATGTAAATGCTGATTA 851

Qy   2572 tatgacccgtcttgttcattaacggtaatttccgttgaacgtaaataacccaataaattt 2631
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    850 TATGACCCATCTTGTTCATTAACGGTAATTTCCGTTGAACGTAAATAACCCAATAAATTT 791

Qy   2632 ttagtcacttttttatagcgagttagcttattaaaggctttagtcaaagctcttaaagac 2691
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    790 TTAGTCACTTTTTTATAACGAGTTAGCTTATTAAAGGCTTTAGTCAAAGCTCTTAAAGAC 731

Qy   2692 acttttaactcctctgctgaatgagcgtttttaacggttaaagttaaaaacaaaaaccgt 2751
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    730 ACTTTTAACTCCTCTGCTGAATGAGCGTTTTTAACGGTTAAAGTTAAAAACAAAAACCGT 671

Qy   2752 cctttaggctctcttgcaactgcttccgcaataatttgttttaactggctcgagtttttc 2811
          |||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||
Db    670 CCTTTAGGCTCTCTTGCAACTGCTTCCGCAATAATTTGTTTTAATTGGCTCGAATTTTTC 611

Qy   2812 atgctccttctccaattacacaatggacacaatcgtttatgacaaaaccacgtttgataa 2871
          |||||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    610 ATGCTCTTTCTCCAATTACACAATGGACATAATCGTTTATGACAAAACCACGTTTGATAA 551

Qy   2872 agttttaagtgctcgccaatcttacgaaaacgcaaaacttcaccacaaccccgtacatca 2931
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    550 AGTTTTAAGTGCTCACCAATCTTACGAAAACGCAAAACTTCACCACAACCCCGTACATCA 491

Qy   2932 tgtgcccgtttaaattctaagattgccaaatattcggcatagcgcacattttcaatcttc 2991
          |||||||||||||||||||||||||||||||||||||||||| || ||||||||||||||
Db    490 TGTGCCCGTTTAAATTCTAAGATTGCCAAATATTCGGCATAACGAACATTTTCAATCTTC 431

Qy   2992 cgttctcgccaaggtctaactttgccattttcagttttatcttcaaaaatttctgacata 3051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 CGTTCTCGCCAAGGTCTAACTTTGCCATTTTCAGTTTTATCTTCAAAAATTTCTGACATA 371

Qy   3052 aaaagctcctccagtttatccacgtgaaggagctgactatc-tttttcaataagcttata 3110
          |||||||||| | | ||  |||||||||||||||||||||| ||||||||||||||||||
Db    370 AAAAGCTCCTTCCGCTTTACCACGTGAAGGAGCTGACTATCTTTTTTCAATAAGCTTATA 311

Qy   3111 accttgacatcatagggctttttcccctagaataggctata-aatcgcaaatgataatcaa 3169
          |||||||||  || |||  ||| | |||| ||||     || ||  | ||| ||||| |||
Db    310 ACCTTGACACCAAAGGCTTTTACACCTAAAATAAAGGCTACAAATGAAAAAGATAAACAA 251

Qy   3170 ctcacgtgttccgagcggccaaactaggaatttgcacgtgggtttttattttgtctttct 3229
          ||||||||  ||      | ||    |  |||| ||||||||||||||||||||||||||||
Db    250 CTCACGTGCTCGGA---CGCCAATCGGAGATTTTCACGTGGGTTTTTATTTTGTCTTTCT 194

Qy   3230 ttcaaccaatttataaccctaataatacaccaaaagcctataaaatcaatggatacaagc 3289
          |||||||||||||||||||||||||||||||  ||| |||||||||||||  | ||||||||
Db    193 TTCAACCAATTTATAACCCTAATAATACACCGAAAACCTATAAAATCAAGGAATACAAGC 134

Qy   3290 ccaattaagcctaatcaagcttgattttaaaaaactagttgttgctaatagtatcaagat 3349
          | ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 CTAATCAAGCCTAATCAAGCTTGATTTTAAAAAACTAGTTGTTGCTAATAGTATCAAGAT 74

Qy   3350 aagaagaaaacgccaaaaatt-gcgtttttaaaccccaaaaagcagatcagcaaaaaccg 3408
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     73 AAGAAGAAAACGCCAAAAATTGGCGTTTTTAAACCCCAAAAAGCAGATCAGCAAAAACCG 14

Qy   3409 ctgaactgctttt 3421
          |||||||||||||
Db     13 CTGAACTGCTTTT 1
```

```
RESULT   8
PCDNA3ZEO
LOCUS       PCDNA3ZEO    3986 bp    DNA             SYN       16-AUG-1995
DEFINITION  Cloning vector pcDNA3ZEO DNA.
ACCESSION   X90639
VERSION     X90639.1  GI:949972
KEYWORDS    cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE      synthetic construct.
  ORGANISM  Plasmid synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 3986)
  AUTHORS   Peters,H., Hundhausen,T., Kroenke,M. and Marget,M.
  TITLE     A new small sized high-level eukaryotic expression vector
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3986)
  AUTHORS   Peters,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
            Michaelisstr.5, D- 24105 Kiel, FRG
COMMENT     Related sequences: M21295 and K03104.
FEATURES             Location/Qualifiers
     source          1. .3986
                     /organism="synthetic construct"
                     /plasmid="pcDNA3ZEO"
                     /db_xref="taxon:32630"
     misc_feature    1. .2125
                     /note="cloning vector (pcDNA3) (Invitrogen)"
     misc_feature    889. .994
                     /note="multiple cloning site (MCS)"
     misc_feature    2126. .2796
                     /note="cloning vector (PZeoSV) (Invitrogen)"
     misc_feature    2797. .3986
                     /note="cloning vector (pcDNA3)"
BASE COUNT      900 a   1057 c   1054 g    975 t
ORIGIN


  Query Match             13.3%;  Score 1082;  DB 56;  Length 3986;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-248;
  Matches 1082;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 GATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCA 269

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA 329

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT 389

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGG 449

Qy    241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC 509

Qy    301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT 569

Qy    361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG 629

Qy    421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC 689

Qy    481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT 749

Qy    541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 AACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA 809

Qy    601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 AGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACG 869

Qy    661 actcactatagggagacccaagcttggtaccgagctcggatccactagtaacggccgcca 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    870 ACTCACTATAGGGAGACCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCA 929

Qy    721 gtgtgctggaattctgcagatatccatcacactggcggccgctcgagcatgcatctagag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    930 GTGTGCTGGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAG 989

Qy    781 ggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgactgtgcct 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    990 GGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCT 1049

Qy    841 tctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctggaaggt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1050 TCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT 1109

Qy    901 gccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtagg 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 GCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGG 1169

Qy    961 tgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaagac 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 TGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC 1229

Qy   1021 aatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccagc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGC 1289

Qy   1081 tg 1082
          ||
Db   1290 TG 1291


RESULT   9
AR098191
LOCUS       AR098191     4026 bp    DNA             PAT       14-FEB-2001
DEFINITION  Sequence 19 from patent US 6074850.
ACCESSION   AR098191
VERSION     AR098191.1  GI:12807448
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4026)
  AUTHORS   Antelman,D., Gregory,R.J. and Wills,K.N.
  TITLE     Retinoblastoma fusion polypeptides
  JOURNAL   Patent: US 6074850-A 19 13-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .4026
                     /organism="unknown"
BASE COUNT      978 a   1020 c    983 g   1045 t
ORIGIN


  Query Match             13.2%;  Score 1068.6;  DB 9;  Length 4026;
  Best Local Similarity   69.7%;  Pred. No. 1.2e-244;
  Matches 1918;  Conservative    0;  Mismatches   89;  Indels  746;  Gaps    5;

Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 GATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCA 269

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA 329

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT 389

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGG 449
```



```
Db   2610 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC 2669

Qy   1715 agagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctg 1774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2670 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTG 2729

Qy   1775 cgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca 1834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2730 CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA 2789

Qy   1835 aaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaa 1894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2790 AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA 2849

Qy   1895 aggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa 1954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2850 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA 2909

Qy   1955 ctcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          |||||||||||||||||||||||||  || | |      |  |  ||||
Db   2910 CTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 2962


RESULT 10
AX060344
LOCUS       AX060344     4597 bp    DNA             PAT       22-JAN-2001
DEFINITION  Sequence 3 from Patent WO0078358.
ACCESSION   AX060344
VERSION     AX060344.1  GI:12405832
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 4597)
  AUTHORS   Chen,W.
  TITLE     Hyaluronic acid microspheres for sustained gene transfer
  JOURNAL   Patent: WO 0078358-A 3 28-DEC-2000;
            The Collaborative Group, Ltd. (US)
FEATURES             Location/Qualifiers
     source          1. .4597
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="pCDNA3.1/GS vector by Invitrogen Corporation"
BASE COUNT     1062 a   1214 c   1206 g   1115 t
ORIGIN


  Query Match           12.2%;  Score 988;  DB 9;  Length 4597;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-225;
  Matches  988;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1076 ccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctc 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3610 CCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC 3669

Qy   1136 ttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatc 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3670 TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC 3729

Qy   1196 agctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaa 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3730 AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA 3789

Qy   1256 catgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtt 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3790 CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT 3849

Qy   1316 tttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtg 1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3850 TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG 3909

Qy   1376 gcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcg 1435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3910 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 3969

Qy   1436 ctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaag 1495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3970 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 4029

Qy   1496 cgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctc 1555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4030 CGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 4089

Qy   1556 caagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaa 1615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4090 CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA 4149

Qy   1616 ctatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg 1675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4150 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 4209

Qy   1676 taacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcc 1735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4210 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC 4269

Qy   1736 taactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttac 1795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4270 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC 4329

Qy   1796 cttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtgg 1855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4330 CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGG 4389

Qy   1856 ttttttttgtttgcaagcagcagattacgcgcagaaaaaaaaggatctcaagaagatccttt 1915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4390 TTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT 4449

Qy   1916 gatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggt 1975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4450 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT 4509

Qy   1976 catgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttccgcgcac 2035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4510 CATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCAC 4569

Qy   2036 atttccccgaaaagtgccacctgacgtc 2063
          |||||||||||||||||||||||||||
Db   4570 ATTTCCCCGAAAAGTGCCACCTGACGTC 4597


RESULT 11
I63120
LOCUS       I63120       4283 bp    DNA             PAT       07-OCT-1997
DEFINITION  Sequence 3 from patent US 5661132.
ACCESSION   I63120
VERSION     I63120.1  GI:2480828
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4283)
  AUTHORS   Eriksson,E., Andree,C., Swain,W.F. and Macklin,M.D.
  TITLE     Wound healing
  JOURNAL   Patent: US 5661132-A 3 26-AUG-1997;
FEATURES             Location/Qualifiers
     source          1. .4283
                     /organism="unknown"
BASE COUNT     1023 a   1095 c   1114 g   1051 t
ORIGIN


  Query Match           12.0%;  Score 973;  DB 10;  Length 4283;
  Best Local Similarity   83.2%;  Pred. No. 9.1e-222;
  Matches 1216;  Conservative    0;  Mismatches   30;  Indels  215;  Gaps    1;

Qy    762 ctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctg 821
```

```
Db   1277 CTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTG 1336

Qy    822 atcagcctcgactgtgccttctagttgccagccatctgttgtttgccccctccccgtgcc 881
Db   1337 ATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCC 1396

Qy    882 ttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgc 941
Db   1397 TTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC 1456

Qy    942 atcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaa 1001
Db   1457 ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA 1516

Qy   1002 gggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttc 1061
Db   1517 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAACC 1576

Qy   1062 tgaggcggaaaga---------------------------------------------- 1074
Db   1577 AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT 1636

Qy   1075 ---------------------------------------------------------- 1074
Db   1637 AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA 1696

Qy   1075 ---------------------------------------------------------- 1074
Db   1697 TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT 1756

Qy   1075 ----------------------------------------------------accagctgcatt 1086
Db   1757 TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATT 1816

Qy   1087 aatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcct 1146
Db   1817 AATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT 1876

Qy   1147 cgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaa 1206
Db   1877 CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 1936

Qy   1207 aggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaa 1266
Db   1937 AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA 1996

Qy   1267 aaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggc 1326
Db   1997 AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCGATAGGC 2056

Qy   1327 tccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccga 1386
Db   2057 TCCGCCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 2116

Qy   1387 caggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttc 1446
Db   2117 CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC 2176

Qy   1447 cgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgcttt 1506
Db   2177 CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 2236

Qy   1507 ctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct 1566
Db   2237 CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT 2296

Qy   1567 gtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg 1626
Db   2297 GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG 2356

Qy   1627 agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatta 1686
```

```
Db   2357 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA 2416

Qy   1687 gcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggct 1746
Db   2417 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCT 2476

Qy   1747 acactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa 1806
Db   2477 ACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 2536

Qy   1807 gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgttt 1866
Db   2537 GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT 2596

Qy   1867 gcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttcta 1926
Db   2597 GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA 2656

Qy   1927 cggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggat 1986
Db   2657 CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 2716

Qy   1987 acatatttgaatgtatttaga 2007
Db   2717 CAAAAAGGATCTTCACCTAGA 2737


RESULT  12
I85496
LOCUS       I85496       4283 bp    DNA             PAT       10-JUN-1998
DEFINITION  Sequence 1 from patent US 5697901.
ACCESSION   I85496
VERSION     I85496.1  GI:3205214
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4283)
  AUTHORS   Eriksson,E.
  TITLE     Gene delivery by microneedle injection
  JOURNAL   Patent: US 5697901-A 1 16-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .4283
                     /organism="unknown"
BASE COUNT     1023 a   1095 c   1114 g   1051 t
ORIGIN


  Query Match           12.0%;  Score 973;  DB 10;  Length 4283;
  Best Local Similarity   83.2%;  Pred. No. 9.1e-222;
  Matches 1216;  Conservative    0;  Mismatches   30;  Indels  215;  Gaps    1;

Qy    762 ctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctg 821
Db   1277 CTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTG 1336

Qy    822 atcagcctcgactgtgccttctagttgccagccatctgttgtttgccccctccccgtgcc 881
Db   1337 ATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCC 1396

Qy    882 ttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgc 941
Db   1397 TTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC 1456

Qy    942 atcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaa 1001
Db   1457 ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA 1516

Qy   1002 gggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttc 1061
Db   1517 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAACC 1576

Qy   1062 tgaggcggaaaga---------------------------------------------- 1074
```

```
Db   1577 AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT 1636

Qy   1075 ------------------------------------------------------------ 1074

Db   1637 AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA 1696

Qy   1075 ------------------------------------------------------------ 1074

Db   1697 TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT 1756

Qy   1075 ----------------------------------------------accagctgcatt 1086

Db   1757 TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATT 1816

Qy   1087 aatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcct 1146

Db   1817 AATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT 1876

Qy   1147 cgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaa 1206

Db   1877 CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 1936

Qy   1207 aggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaa 1266

Db   1937 AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA 1996

Qy   1267 aaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggc 1326

Db   1997 AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCGATAGGC 2056

Qy   1327 tccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccga 1386

Db   2057 TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 2116

Qy   1387 caggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttc 1446

Db   2117 CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC 2176

Qy   1447 cgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgcttt 1506

Db   2177 CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 2236

Qy   1507 ctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct 1566

Db   2237 CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT 2296

Qy   1567 gtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg 1626

Db   2297 GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG 2356

Qy   1627 agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatta 1686

Db   2357 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA 2416

Qy   1687 gcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggct 1746

Db   2417 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCT 2476

Qy   1747 acactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa 1806

Db   2477 ACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 2536

Qy   1807 gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgttt 1866

Db   2537 GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT 2596

Qy   1867 gcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttcta 1926

Db   2597 GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA 2656

Qy   1927 cggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggat 1986

Db   2657 CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 2716

Qy   1987 acatatttgaatgtatttaga 2007

Db   2717 CAAAAAGGATCTTCACCTAGA 2737
```

```
RESULT 13
AF053407
LOCUS       AF053407     4050 bp    DNA   circular  SYN       17-OCT-1998
DEFINITION  Expression vector pMAS-S, complete sequence.
ACCESSION   AF053407
VERSION     AF053407.1  GI:3746483
KEYWORDS    .
SOURCE      Expression vector pMAS-S.
  ORGANISM  Expression vector pMAS-S
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4050)
  AUTHORS   Krieg,A.M., Wu,T., Weeratna,R., Efler,S.M., Love-Homan,L., Yang,L.,
            Yi,A.K., Short,D. and Davis,H.L.
  TITLE     Sequence motifs in adenoviral DNA block immune activation by
            stimulatory CpG motifs
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)
  MEDLINE   98445422
REFERENCE   2  (bases 1 to 4050)
  AUTHORS   Wu,T., Efler,S.M., Davis,H.L., Krieg,A.M. and Schorr,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-1998) HGD, Loeb Health Research Institute, 725
            Parkdale Ave., Ottawa, ON K1Y 4E9, Canada
FEATURES             Location/Qualifiers
     source          1. .4050
                     /organism="Expression vector pMAS-S"
                     /db_xref="taxon:82305"
     gene            3351. .4031
                     /gene="S"
     CDS             3351. .4031
                     /gene="S"
                     /note="derived from Hepatitis B virus surface antigen"
                     /codon_start=1
                     /transl_table=11
                     /product="surface antigen subtype ayw"
                     /protein_id="AAC64340.1"
                     /db_xref="GI:3746484"
                     /translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGG
                     TTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQG
                     MLPVCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGK
                     FLWEWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPI
                     FFCLWVYI"
BASE COUNT      993 a   1024 c    949 g   1084 t
ORIGIN
```

```
  Query Match            11.9%;  Score 968.6;  DB 56;  Length 4050;
  Best Local Similarity   85.9%;  Pred. No. 1e-220;
  Matches 1155;  Conservative    0;  Mismatches   34;  Indels  156;  Gaps    1;

Qy    796 tgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagcca 855

Db     40 TCTCGAGTCTAGACTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA 99

Qy    856 tctgttgtttgcccctccccgtgccttccttgaccctggaaggtgccactcccactgtc 915

Db    100 TCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTC 159

Qy    916 ctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctg 975

Db    160 CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG 219

Qy    976 gggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgct 1035

Db    220 GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT 279

Qy   1036 ggggatgcggtgggctctatggc---------------------------------- 1058
```

```
Db    280 GGGGAAGGCCTCGGACTAGTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT 339

Qy   1059 ------------------------------------------------------------ 1058

Db    340 TATCCGCTCACAATTCCACACAACATACGAGCCGCGGAAGCATAAAGTGTAAAGCCTGGG 399

Qy   1059 ------------------------------------------------------------t 1059

Db    400 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 459

Qy   1060 tctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119

Db    460 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 519

Qy   1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179

Db    520 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 579

Qy   1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239

Db    580 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 639

Qy   1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299

Db    640 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 699

Qy   1300 ccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgac 1359

Db    700 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 759

Qy   1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 1419

Db    760 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 819

Qy   1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479

Db    820 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 879

Qy   1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539

Db    880 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 939

Qy   1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgct 1599

Db    940 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 999

Qy   1600 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659

Db   1000 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 1059

Qy   1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719

Db   1060 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 1119

Qy   1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779

Db   1120 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC 1179

Qy   1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839

Db   1180 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 1239

Qy   1840 ccgctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899

Db   1240 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 1299

Qy   1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959

Db   1300 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 1359

Qy   1960 gttaagggattttggtcatgagcgg 1984
```

```
Db   1360 GTTAAGGGATTTTGGTCATGAGCTG 1384


RESULT 14
AF053408
LOCUS       AF053408     4227 bp    DNA   circular  SYN       17-OCT-1998
DEFINITION  Expression vector pMCG16-S, complete sequence.
ACCESSION   AF053408
VERSION     AF053408.1  GI:3746485
KEYWORDS    .
SOURCE      Expression vector pMCG16-S.
  ORGANISM  Expression vector pMCG16-S
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4227)
  AUTHORS   Krieg,A.M., Wu,T., Weeratna,R., Efler,S.M., Love-Homan,L., Yang,L.,
            Yi,A.K., Short,D. and Davis,H.L.
  TITLE     Sequence motifs in adenoviral DNA block immune activation by
            stimulatory CpG motifs
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)
  MEDLINE   98445422
REFERENCE   2  (bases 1 to 4227)
  AUTHORS   Wu,T., Efler,S.M., Davis,H.L., Krieg,A.M. and Schorr,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-1998) HGD, Loeb Health Research Institute, 725
            Parkdale Ave., Ottawa, ON K1Y 4E9, Canada
FEATURES             Location/Qualifiers
     source          1. .4227
                     /organism="Expression vector pMCG16-S"
                     /db_xref="taxon:82306"
     gene            3528. .4208
                     /gene="S"
     CDS             3528. .4208
                     /gene="S"
                     /note="derived from Hepatitis B virus surface antigen"
                     /codon_start=1
                     /transl_table=11
                     /product="surface antigen subtype ayw"
                     /protein_id="AAC64341.1"
                     /db_xref="GI:3746486"
                     /translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGG
                     TTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQG
                     MLPVCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGK
                     FLWEWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPI
                     FFCLWVYI"
BASE COUNT     1021 a   1076 c    990 g   1140 t
ORIGIN


  Query Match           11.9%;  Score 968.6;  DB 56;  Length 4227;
  Best Local Similarity  85.9%;  Pred. No. 1e-220;
  Matches 1155;  Conservative    0;  Mismatches   34;  Indels  156;  Gaps    1;

Qy    796 tgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagcca 855

Db     40 TCTCGAGTCTAGACTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA 99

Qy    856 tctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtc 915

Db    100 TCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTC 159

Qy    916 ctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctg 975

Db    160 CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG 219

Qy    976 gggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgct 1035

Db    220 GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT 279

Qy   1036 ggggatgcggtgggctctatggc---------------------------------------- 1058

Db    280 GGGGAAGGCCTCGGACTAGTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT 339

Qy   1059 ------------------------------------------------------------ 1058
```

```
Db    340 TATCCGCTCACAATTCCACACAACATACGAGCCGCGGAAGCATAAAGTGTAAAGCCTGGG 399

Qy   1059 ------------------------------------------------------------t 1059
                                                                        |
Db    400 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 459

Qy   1060 tctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119
             ||  |      |  ||||||||||||||||||||||||||||||||||||||||||
Db    460 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 519

Qy   1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    520 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 579

Qy   1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    580 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 639

Qy   1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    640 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 699

Qy   1300 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    700 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 759

Qy   1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    760 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 819

Qy   1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    820 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 879

Qy   1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539
          ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db    880 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 939

Qy   1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccccgttcagcccgaccgct 1599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    940 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 999

Qy   1600 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1000 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 1059

Qy   1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 1119

Qy   1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779
          ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   1120 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC 1179

Qy   1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1180 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 1239

Qy   1840 ccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1240 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 1299

Qy   1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1300 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 1359

Qy   1960 gttaagggattttggtcatgagcgg 1984
          ||||||||||||||||||||||| |
Db   1360 GTTAAGGGATTTTGGTCATGAGCTG 1384


RESULT 15
```

```
AF053409
LOCUS       AF053409     4625 bp    DNA   circular  SYN       17-OCT-1998
DEFINITION  Expression vector pMCG50-S, complete sequence.
ACCESSION   AF053409
VERSION     AF053409.1  GI:3746487
KEYWORDS    .
SOURCE      Expression vector pMCG50-S.
  ORGANISM  Expression vector pMCG50-S
            artificial sequence; vectors.
REFERENCE   1  (bases 1 to 4625)
  AUTHORS   Krieg,A.M., Wu,T., Weeratna,R., Efler,S.M., Love-Homan,L., Yang,L.,
            Yi,A.K., Short,D. and Davis,H.L.
  TITLE     Sequence motifs in adenoviral DNA block immune activation by
            stimulatory CpG motifs
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)
  MEDLINE   98445422
REFERENCE   2  (bases 1 to 4625)
  AUTHORS   Wu,T., Efler,S.M., Davis,H.L., Krieg,A.M. and Schorr,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAR-1998) HGD, Loeb Health Research Institute, 725
            Parkdale Ave., Ottawa, ON K1Y 4E9, Canada
FEATURES             Location/Qualifiers
     source          1. .4625
                     /organism="Expression vector pMCG50-S"
                     /db_xref="taxon:82307"
     gene            3926. .4606
                     /gene="S"
     CDS             3926. .4606
                     /gene="S"
                     /note="derived from Hepatitis B virus surface antigen"
                     /codon_start=1
                     /transl_table=11
                     /product="surface antigen subtype ayw"
                     /protein_id="AAC64342.1"
                     /db_xref="GI:3746488"
                     /translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGG
                     TTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQG
                     MLPVCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGK
                     FLWEWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPI
                     FFCLWVYI"
BASE COUNT     1084 a   1193 c   1082 g   1266 t
ORIGIN


  Query Match             11.9%;  Score 968.6;  DB 56;  Length 4625;
  Best Local Similarity   85.9%;  Pred. No. 1e-220;
  Matches 1155;  Conservative    0;  Mismatches   34;  Indels  156;  Gaps    1;

Qy    796 tgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagcca 855
          | ||   |  |  ||||||||||||||||||||||||||||||||||||||||||||||
Db     40 TCTCGAGTCTAGACTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCA 99

Qy    856 tctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgtc 915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 TCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTC 159

Qy    916 ctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctg 975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 CTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTG 219

Qy    976 gggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgct 1035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 GGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCT 279

Qy   1036 ggggatgcggtgggctctatggc---------------------------------- 1058
          |||||  |   | ||     ||||
Db    280 GGGGAAGGCCTCGGACTAGTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGT 339

Qy   1059 ------------------------------------------------------------ 1058

Db    340 TATCCGCTCACAATTCCACACAACATACGAGCCGCGGAAGCATAAAGTGTAAAGCCTGGG 399

Qy   1059 ------------------------------------------------------------t 1059
```

```
Db    400 GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT 459

Qy   1060 tctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119
          ||  |     |  ||||||||||||||||||||||||||||||||||||||||||||
Db    460 CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT 519

Qy   1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    520 TGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 579

Qy   1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    580 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 639

Qy   1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    640 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 699

Qy   1300 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    700 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 759

Qy   1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    760 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 819

Qy   1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    820 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 879

Qy   1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539
          ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db    880 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 939

Qy   1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgct 1599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    940 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 999

Qy   1600 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1000 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 1059

Qy   1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 1119

Qy   1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779
          ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   1120 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC 1179

Qy   1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1180 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 1239

Qy   1840 ccgctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1240 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 1299

Qy   1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1300 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 1359

Qy   1960 gttaagggattttggtcatgagcgg 1984
          |||||||||||||||||||||| |
Db   1360 GTTAAGGGATTTTGGTCATGAGCTG 1384
```

Search completed: November 3, 2001, 20:31:24
Job time: 17035 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 16:01:04 ; Search time 444.17 Seconds
(without alignments)
11471.782 Million cell updates/sec

Title: US-09-778-516-1
Perfect score: 8115
Sequence: 1 gatgtacgggccagatatac.....gcgtgtcgctagcgacggtc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	3497.2	43.1	5059	21	AAZ36227	DNA encoding a per
	2	3110.6	38.3	3203	11	AAQ06760	Sequence of beta-g
	3	1245.2	15.3	3853	19	AAV40006	Plasmid pCTM. Chi
	4	1152.4	14.2	5306	21	AAA53878	Expression vector
c	5	1148.2	14.1	14455	19	AAV32374	Complete sequence
c	6	1148.2	14.1	14455	21	AAA59050	Nucleotide sequenc
	7	1140	14.0	5377	21	AAA53872	Expression vector
	8	1131.4	13.9	4645	21	AAA53870	Expression vector
	9	1082	13.3	5446	19	AAV38297	Plasmid pcDNA3. C
	10	1071.6	13.2	5245	21	AAA53871	Expression vector
	11	1070.2	13.2	4026	19	AAV40007	Plasmid pCTMI. Ch
	12	1068	13.2	8705	20	AAZ23778	Vector pShuttle DN
	13	1058.4	13.0	6050	21	AAA47797	pWRG/SEO-S vaccine
	14	1058.4	13.0	8002	21	AAA47796	pWRG/SEO-M vaccine
	15	1036.2	12.8	3796	21	AAA27831	Vector plasmid pCM
	16	988	12.2	4597	22	AAF24901	Nucleotide sequenc
	17	984.6	12.1	8136	18	AAT63236	Plasmid pCR(TM)3la
	18	981.8	12.1	6253	20	AAX08454	AAV vector sequenc
	19	981.8	12.1	6280	20	AAX08455	AAV vector sequenc
	20	981.8	12.1	6280	20	AAX08456	AAV vector sequenc
	21	981.8	12.1	6565	21	AAZ45925	Nucleotide sequenc
	22	981.8	12.1	6981	21	AAZ45931	Nucleotide sequenc
	23	981.8	12.1	7054	21	AAZ45927	Nucleotide sequenc
	24	981.8	12.1	7405	21	AAZ45926	Nucleotide sequenc
	25	973	12.0	4283	18	AAT86449	DNA encoding hGH,
	26	973	12.0	4283	19	AAT86610	Epidermal growth f
	27	973	12.0	4283	21	AAA63237	Plasmid pWRG1630 n
	28	961.4	11.8	6200	21	AAZ49493	Human alpha-L-idur
	29	959.6	11.8	6115	18	AAT63235	Eukaryotic express
	30	948.4	11.7	5857	21	AAA95423	Expression vector
	31	935.6	11.5	5432	21	AAZ89476	Transgenic APP DNA
	32	932.8	11.5	5732	21	AAA71429	Expression vector
	33	932	11.5	5729	21	AAA71430	Expression vector
	34	921.8	11.4	6281	22	AAC84206	Plasmid Glambda-1A
	35	921.8	11.4	6284	19	AAV41427	Plasmid Hul9AHcpcd
	36	921.8	11.4	6285	17	AAT34109	Anti-IL-5 humanise
	37	921.8	11.4	6285	19	AAV03503	Plasmid pCDIL5HZHC
	38	912.6	11.2	6094	21	AAA71428	Expression vector
c	39	908.2	11.2	5957	21	AAC55464	Destination vector
	40	908.2	11.2	5957	21	AAC55467	Destination vector
	41	906.4	11.2	2704	20	AAZ06758	Vector pUC28 nucle
	42	906.4	11.2	3130	19	AAV46334	Universal reporter
c	43	906.4	11.2	3344	21	AAA39495	Transgenic unc-119
	44	906.4	11.2	3789	21	AAZ39782	Plasmid pGB1MMN nu
c	45	906.4	11.2	4173	19	AAV32663	Plasmid pGFP-TT re

ALIGNMENTS

RESULT 1
AAZ36227
ID AAZ36227 standard; DNA; 5059 BP.
XX
AC AAZ36227;
XX
DT 22-FEB-2000 (first entry)
XX
DE DNA encoding a permease an beta-galactosidase.
XX
KW Lactose operon; permease; beta-galactosidase; mutant; yoghurt;
KW beta-galactosidase activity; lactose fermentation;
KW fermented milk product; Streptococcus thermophilus; ds.
XX
OS Lactobacillus bulgaricus.
XX
FH Key Location/Qualifiers
FT CDS 122..1873
FT /*tag= a
FT /product= "permease"
FT /note= "encodes AAY53620"
FT CDS 1877..4519
FT /*tag= b
FT /product= "beta-galactosidase"
FT /note= "encodes AAY53621"
XX
PN FR2778921-A1.
XX
PD 26-NOV-1999.
XX
PF 22-MAY-1998; 98FR-0006456.
XX
PR 22-MAY-1998; 98FR-0006456.
XX

PA (DANO-) CIE DANONE SA GERVAIS.
XX
PI Benbadis L, Brignon P, Gendre F;
XX
DR WPI; 2000-026146/03.
DR P-PSDB; AAY53620, AAY53621.
XX
PT New mutant Lactobacillus bulgaricus strain lacking beta-galactosidase
PT activity useful for making fermented milk products, especially yoghurt
PT -
XX
PS Disclosure; Page 15-22; 32pp; French.
XX
CC The present sequence represents a lactose operon of Lactobacillus
CC bulgaricus, and encodes a permease protein and beta-galactosidase
CC protein. The specification describes mutant L. bulgaricus strain lacking
CC beta-galactosidase activity, because it has a termination (nonsense)
CC mutation in at least one of the coding sequences of the lactose operon.
CC The resulting mutant strain is unable to ferment lactose. Since the
CC mutant strain cannot ferment lactose, fermented milk products produced
CC using it do not have to be cooled to prevent post-acidification. The
CC mutant L. bulgaricus strain of the invention is used for producing a
CC fermented milk product, especially yoghurt. Especially, milk is
CC fermented with a culture of the mutant strain and optionally
CC Streptococcus thermophilus in the presence of at least one sugar
CC assimilable by the mutant strain, preferably glucose.
XX
SQ Sequence 5059 BP; 1229 A; 1341 C; 1323 G; 1166 T; 0 other;

Query Match 43.1%; Score 3497.2; DB 21; Length 5059;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3518; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

```
Qy  4584 ccgataaaaagttaggcgacggctttgccctggtgccagcagacggtaaggtctacgcgc 4643
         |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1479 ccgacaaaaagttaggcgacggctttgccctggtgccagcagacggtaaggtctacgcgc 1538

Qy  4644 catttgccggtactgtccgccagctggccaagacccggcactcgatcgtcctggaaaatg 4703
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1539 catttgccggtactgtccgccagctggccaagacccggcactcgatcgtcctggaaaatg 1598

Qy  4704 aacatggggtcttggtcttgattcaccttggcctgggcacggtcaaattaaacgggactg 4763
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1599 aacatggggtcttggtcttgattcaccttggcctgggcacggtcaaattaaacgggactg 1658

Qy  4764 gctttgtcagctatgttgaagagggcagccaggtagaagccggccagcagatcctggaat 4823
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1659 gctttgtcagctatgttgaagagggcagccaggtagaagccggccagcagatcctggaat 1718

Qy  4824 tctgggacccggcgatcaagcaggccaagctggacgacacggtaatcgtgaccgtcatca 4883
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1719 tctgggacccggcgatcaagcaggccaagctggacgacacggtaatcgtgaccgtcatca 1778

Qy  4884 acagcgaaactttcacaaatagccagatgctcttgccgatcggccacagcgtccaagccc 4943
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  1779 acagcgaaactttcgcaaatagccagatgctcttgccgatcggccacagcgtccaagccc 1838

Qy  4944 tggatgatgtattcaagttagaagggaagaattagaaaatgagcaataagttagtaaaag 5003
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1839 tggatgatgtattcaagttagaagggaagaattagaaaatgagcaataagttagtaaaag 1898

Qy  5004 aaaaaagagttgaccaggcagacctggcctggctgactgacccggaagtttacgaagtca 5063
         |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  1899 aaaaaagagttgaccaggcagacttggcctggctgactgacccggaagtttacgaagtca 1958

Qy  5064 atacaattccccccgcactccgaccatgagtccttccaaagccaggaagaactggaggagg 5123
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1959 atacaattccccccgcactccgaccatgagtccttccaaagccaggaagaactggaggagg 2018

Qy  5124 gcaagtccagtttagtgcagtccctggacggggactggctgattgactacgctgaaaacg 5183
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2019 gcaagtccagtttagtgcagtccctggacggggactggctgattgactacgctgaaaacg 2078

Qy  5184 gccagggaccagtcaacttctatgcagaagactttgacgatagcaattttaagtcagtca 5243
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2079 gccagggaccagtcaacttctatgcagaagactttgacgatagcaattttaagtcagtca 2138

Qy  5244 aagtacccggcaacctggaactgcaaggctttggccagccccagtatgtcaacgtccaat 5303
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2139 aagtacccggcaacctggaactgcaaggctttggccagccccagtatgtcaacgtccaat 2198

Qy  5304 atccatgggacggcagtgaggagattttcccgccccaaattccaagcaaaaatccgctcg 5363
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2199 atccatgggacggcagtgaggagattttcccgccccaaattccaagcaaaaatccgctcg 2258

Qy  5364 cttcttatgtcagatactttgacctggatgaagctttctgggacaaggaagtcagcttga 5423
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2259 cttcttatgtcagatactttgacctggatgaagctttctgggacaaggaagtcagcttga 2318

Qy  5424 agtttgacggggcggcaacagccatctatgtctggctgaacggccacttcgtcggctacg 5483
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2319 agtttgacggggcggcaacagccatctatgtctggctgaacggccacttcgtcggctacg 2378

Qy  5484 gggaagactcctttaccccaagcgagtttatggttaccaagttcctcaagaaagaaaata 5543
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2379 gggaagactcctttaccccaagcgagtttatggttaccaagttcctcaagaaagaaaata 2438

Qy  5544 accgcctggcagtggctctctacaagtattcttccgcctcctggctggaagaccaggact 5603
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2439 accgcctggcagtggctctctacaagtattcttccgcctcctggctggaagaccaggact 2498

Qy  5604 tctggcgcatgtctggtttgttcagatcagtgactcttcaggccaagccgcgtctgcact 5663
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2499 tctggcgcatgtctggtttgttcagatcagtgactcttcaggccaagccgcgtctgcact 2558

Qy  5664 tggaggaccttaagcttacggccagcttgaccgataactaccaaaaaggaaagctggaag 5723
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2559 tggaggaccttaagcttacggccagcttgaccgataactaccaaaaaggaaagctggaag 2618

Qy  5724 tcgaagccaatattgcctaccgcttgccaaatgccagctttaagctggaagtgcgggata 5783
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2619 tcgaagccaatattgcctaccgcttgccaaatgccagctttaagctggaagtgcgggata 2678

Qy  5784 gtgaaggtgacttggttgctgaaaagctgggcccaatcagaagcgagcagctggaattca 5843
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2679 gtgaaggtgacttggttgctgaaaagctgggcccaatcagaagcgagcagctggaattca 2738

Qy  5844 ctctggctgatttgccagtagctgcctggagcgcggaaaagcctaacctttaccaggtcc 5903
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2739 ctctggctgatttgccagtagctgcctggagcgcggaaaagcctaacctttaccaggtcc 2798

Qy  5904 gcctgtatttataccaggcaggcagcctcttagaggttagccggcaggaagtgggtttcc 5963
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2799 gcctgtatttataccaggcaggcagcctcttagaggttagccggcaggaagtgggtttcc 2858

Qy  5964 gcaactttgaactaaaagacgggattatgtaccttaacggccagcggatcgtcttcaagg 6023
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2859 gcaactttgaactaaaagacgggattatgtaccttaacggccagcggatcgtcttcaagg 2918

Qy  6024 gggccaaccggcacgaatttgacagtaagttgggtcgggctatcacggaagaggatatga 6083
         |||||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||
Db  2919 gggccaaccggcacgaatttgacagtaagttgggccgggctatcacagaagaggatatga 2978

Qy  6084 tctgggacatcaagaccatgaagcgaagcaacatcaatgctgtccgctgctctcactacc 6143
         ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2979 tctgggatatcaagaccatgaagcgaagcaacatcaatgctgtccgctgctctcactacc 3038

Qy  6144 cgaaccagtccctctttaccggctctgtgacaagtacggcctttacgtcattgatgaag 6203
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3039 cgaaccagtccctctttaccggctctgtgacaagtacggcctttacgtcattgatgaag 3098

Qy  6204 ctaacctggaaagccacggcacctgggaaaaagtgggggggcacgaagatcctagcttca 6263
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3099 ctaacctggaaagccacggcacctgggaaaaagtgggggggcacgaagatcctagcttca 3158
```

```
Qy   6264 atgttccaggcgatgaccagcattggctgggagccagcttatcccgggtgaagaacatga 6323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3159 atgttccaggcgatgaccagcattggctgggagccagcttatcccgggtgaagaacatga 3218

Qy   6324 tggctcgggacaagaaccatgcttcaatcctaatctggtctttaggcaatgagtcttacg 6383
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   3219 tggctcgggacaagaaccatgcttcaatcctgatctggtctttaggcaatgagtcttacg 3278

Qy   6384 ccggcactgtctttgcccaaatggctgattacgtccggaaggctgatccgacccgggttc 6443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3279 ccggcactgtctttgcccaaatggctgattacgtccggaaggctgatccgacccgggttc 3338

Qy   6444 agcactatgaaggggtgacccacaaccggaagtttgacgacgccacccagattgaaagcc 6503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3339 agcactatgaaggggtgacccacaaccggaagtttgacgacgccacccagattgaaagcc 3398

Qy   6504 ggatgtatgctccggccaaggtaattgaagaatacttgaccaataaaccagccaagccat 6563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3399 ggatgtatgctccggccaaggtaattgaagaatacttgaccaataaaccagccaagccat 3458

Qy   6564 ttatctcagttgaatacgctcacgccatgggcaactccgtcggtgacctggccgcctaca 6623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3459 ttatctcagttgaatacgctcacgccatgggcaactccgtcggtgacctggccgcctaca 3518

Qy   6624 cggccctggaaaaatacccccactaccagggcggcttcatctgggactggattgaccaag 6683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3519 cggccctggaaaaatacccccactaccagggcggcttcatctgggactggattgaccaag 3578

Qy   6684 gactggaaaaagacgggcacctgctttatgggggcgacttcgatgaccggccaaccgact 6743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3579 gactggaaaaagacgggcacctgctttatgggggcgacttcgatgaccggccaaccgact 3638

Qy   6744 atgaattctgcgggaacggcctggtctttgctgaccggactgaatcgccgaaactggcta 6803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3639 atgaattctgcgggaacggcctggtctttgctgaccggactgaatcgccgaaactggcta 3698

Qy   6804 atgtcaaggccctttacgccaaccttaagttagaagtaaaagatgggcagctcttcctca 6863
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3699 atgtcaaggccctttacgccaaccttaagttagaagtaaaagatgggcagctcttcctca 3758

Qy   6864 aaaacgacaatttatttaccaacagctcatcttactacttcttgactagtcttttggtcg 6923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3759 aaaacgacaatttatttaccaacagctcatcttactacttcttgactagtcttttggtcg 3818

Qy   6924 atggcaagttgacctaccagagccggcctctgacctttggcctggagcctggcgaatccg 6983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3819 atggcaagttgacctaccagagccggcctctgacctttggcctggagcctggcgaatccg 3878

Qy   6984 ggacctttgccctgccttggccggaagtcgctgatgaaaaaggggaggtcgtctaccggg 7043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3879 ggacctttgccctgccttggccggaagtcgctgatgaaaaaggagaggtcgtctaccggg 3938

Qy   7044 taacggcccacttaaaagaagacttgccttgggcggatgagggcttcactgtggctgaag 7103
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3939 taacggcccacttaaaagaagacttgccttgggcggatgagggcttcactgtggctgaag 3998

Qy   7104 cagaagaagtagctcaaaagctgccggaatttaagccggaagggcggccagatttagttg 7163
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3999 cagaagaagtagctcaaaagctgccggaatttaagccggaagggcggccagatttagttg 4058

Qy   7164 attccgactacaacctaggcctgaaaggaaataacttccaaattctcttctccaaggtca 7223
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4059 attccgactacaacctaggcctgaaaggaaataacttccaaattctcttctccaaggtca 4118

Qy   7224 agggctggccggtttccctcaagtatgccggtagggaatacttgaagcggctgccggaat 7283
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4119 agggctggccggtttccctcaagtatgccggtagggaatacttgaagcggctgccggaat 4178

Qy   7284 ttaccttctggcgggccctgacggacaacgaccggggagctggttacggctatgatctgg 7343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4179 ttaccttctggcgggccctgacggacaacgaccggggagctggttacggctatgatctgg 4238

Qy   7344 cccggtgggaaaatgccggcaagtatgcccgcttgaaagacatcagctgcgaggtcaagg 7403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4239 cccggtgggaaaatgccggcaagtatgcccgcttgaaagacatcagctgcgaggtcaagg 4298

Qy   7404 aagactccgttttggtcaagactgcctttacgttgcctgtcgccttaaagggtgatttaa 7463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4299 aagactccgttttggtcaagactgcctttacgttgcctgtcgccttaaagggtgatttaa 4358

Qy   7464 ccgtgacctatgaagtcgatggacggggcaagattgctgtaacagctgacttcccaggcg 7523
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4359 ctgtgacctatgaagtcgatggacggggcaagattgctgtaacagctgacttcccaggcg 4418

Qy   7524 cggaagaagctggtctcttgccagcctttggcttgaacctggccctgccaaaagaactga 7583
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4419 cggaagaagccggtctcttgccagcctttggcttgaacctggccctgccaaaagaactga 4478

Qy   7584 ccgattaccgctactatggtctgggacctaatgagagctacccagaccgcttggaaggta 7643
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   4479 ccgattaccgctactatggtctgggacctaatgagagctaaccagaccgcttggaaggta 4538

Qy   7644 attacctgggcatctaccagggagcggtaaaaaagaactttagcccata---tcgtccgc 7700
          |||||||||||||||||||||||||||||||||||||||||||||||    |||||||
Db   4539 attacctgggcatctaccagggagcggtaaaaaagaactttagcccatacctgcgtccgc 4598

Qy   7701 aggaaacgggcaaccggagcaaggttcgctggtaccagctctttgatgaaaagggcggct 7760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4599 aggaaacgggcaaccggagcaaggttcgctggtaccagctctttgatgaaaagggcggct 4658

Qy   7761 tggaatttacggccaatggggcagacttgaacttgtctgctttgccatattctgccgccc 7820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4659 tggaatttacggccaatggggcagacttgaacttgtctgctttgccatattctgccgccc 4718

Qy   7821 aaattgaagcagcggaccacgcttttgaactgactaacaattacacttgggttagagcct 7880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4719 aaattgaagcagcggaccacgcttttgaactgactaacaattacacttgggttagagcct 4778

Qy   7881 taagcgcccagatgggggtcggcggggatgactcctgggggcagaaggtccacccggaat 7940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4779 taagcgcccagatgggggtcggcggggatgactcctgggggcagaaggtccacccggaat 4838

Qy   7941 tctgcctggatgctcaaaaagcccgccagcttcgcctggtgattcagccccttttactaa 8000
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   4839 tctgcctggatgctcaaaaagcccgccagctccgcctggtgattcagccccttttactaa 4898

Qy   8001 aataaatgctacaattgacttaacaggatgaaattttagtaaaagcaaagcgagtgagga 8060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4899 aataaatgctacaattgacttaacaggatgaaattttagtaaaagcaaagcgagtgagga 4958

Qy   8061 agatggcaacgatcagagaagtgccaaggcagccggcgtgtcgctagcgacggt 8114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4959 agatggcaacgatcagagaagtgccaaggcagccggcgtgtcgctagcgacggt 5012
```

```
RESULT   2
AAQ06760
ID   AAQ06760 standard; DNA; 3203 BP.
XX
AC   AAQ06760;
XX
DT   25-FEB-1991  (first entry)
XX
DE   Sequence of beta-galactosidase gene.
XX
KW   pBR329; pBG1; ds.
XX
OS   Lactobacillus bulgaricus SBT0034.
XX
FH   Key             Location/Qualifiers
FT   CDS             285..3083
FT                   /*tag= a
XX
PN   JP02261383-A.
XX
PD   24-OCT-1990.
```

```
XX
PF   31-MAR-1989;   89JP-0082851.
XX
PR   31-MAR-1989;   89JP-0082851.
XX
PA   (SNOW ) SNOW BRAND MILK PRODUCTS.
XX
DR   WPI; 1990-364398/49.
DR   P-PSDB; AAR08132.
XX
PT   DNA from Lactobacillus bulgaricus STB0034 encoding beta-
PT   galactosidase - used to produce high yields of beta-galactosidase by
PT   recombinant DNA technology.
XX
PS   Claim 1; Fig 1; 6pp; Japanese.
XX
CC   Beta-galactosidase may be produced from an E.coli expression system
CC   transformed by the gene sequence incorporated in the plasmid vector
CC   pBG1, derived from pBR329.
XX
SQ   Sequence 3203 BP; 836 A; 807 C; 867 G; 693 T; 0 other;


  Query Match             38.3%;  Score 3110.6;  DB 11;  Length 3203;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 3177;  Conservative    0;  Mismatches   14;  Indels    8;  Gaps    5;

Qy   4920 cgatcggccacagcgtccaagccctggatgatgtattcaagttagaagggaagaattaga 4979
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 cgatcggccacagcgtccaagccctggatgatgtattcaagttagaagggaagaattaga 60

Qy   4980 aaatgagcaataagttagtaaaagaaaaaagagttgaccaggcagacctggcctggctga 5039
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 aaatgagcaataagttagtaaaagaaaaaagagttgaccaggcagacctggcctggctga 120

Qy   5040 ctgacccggaagtttacgaagtcaatacaattcccccgcactccgaccatgagtccttcc 5099
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ctgacccggaagtttacgaagtcaatacaattcccccgcactccgaccatgagtccttcc 180

Qy   5100 aaagccaggaagaactggaggagggcaagtccagtttagtgcagtccctggacggggact 5159
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    181 aaagccaggaagaactggaggagggcaagtccagtttagtgcagtccctggac-gggact 239

Qy   5160 ggctgattgactacgctgaaaacggccagggaccagtcaacttctatgcagaagactttg 5219
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ggctgattgactacgctgaaaacggccagggaccagtcaacttctatgcagaagactttg 299

Qy   5220 acgatagcaattttaagtcagtcaaagtacccggcaacctggaactgcaaggctttggcc 5279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 acgatagcaattttaagtcagtcaaagtacccggcaacctggaactgcaaggctttggcc 359

Qy   5280 agccccagtatgtcaacgtccaatatccatgggacggcagtgaggagattttcccgcccc 5339
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 agccccagtatgtcaacgtccaatatccatgggacggcagtgaggagattttcccgcccc 419

Qy   5340 aaattccaagcaaaaatccgctcgcttcttatgtcagatactttgacctggatgaagctt 5399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 aaattccaagcaaaaatccgctcgcttcttatgtcagatactttgacctggatgaagctt 479

Qy   5400 tct-gggacaaggaagtcagcttgaagtttgacggggcggcaacagccatctatgtctgg 5458
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 tctggggacaaggaagtcagcttgaagtttgacggggcggcaacagccatctatgtctgg 539

Qy   5459 ctgaacggccacttcgtcggctacggggaagactcctttaccccaagcgagtttatggtt 5518
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    540 ctgaacggccacttcgtcggctacggggaagactcctttaccccaagtgagtttatggtt 599

Qy   5519 accaagttcctcaagaaagaaaataaccgcctggcagtggctctctacaagtattcttcc 5578
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 accaagttcctcaagaaagaaaataaccgcctggcagtggctctctacaagtattcttcc 659

Qy   5579 gcctcctggctggaagaccaggacttctggcgcatgtctggtttgttcagatcagtgact 5638
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    660 gcctcctggctggaagaccaggacttctggcgcatgtctggtttgttcagatcagtgact 719

Qy   5639 cttcaggccaagccgcgtctgcacttggaggaccttaagcttacggccagcttgaccgat 5698
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    720 cttcaggccaagccgcgtctgcacttggaggaccttaagcttacggccagcttgaccgat 779

Qy   5699 aactaccaaaaaggaaagctggaagtcgaagccaatattgcctaccgcttgccaaatgcc 5758
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    780 aactaccaaaaaggaaagctggaagtcgaagccaatattgcctaccgcttgccaaatgcc 839

Qy   5759 agctttaagctggaagtgcgggatagtgaaggtgacttggttgctgaaaagctgggccca 5818
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    840 agctttaagctggaagtgcgggatagtgaaggtgacttggttgctgaaaagctgggccca 899

Qy   5819 atcagaagcgagcagctggaattcactctggctgatttgccagtagctgcctggagcgcg 5878
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 atcagaagcgagcagctggaattcactctggctgatttgccagtagctgcctggagcgcg 959

Qy   5879 gaaaagcctaacctttaccaggtccgcctgtatttataccaggcaggcagcctcttagag 5938
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    960 gaaaagcctaacctttaccaggtccgcctgtatttataccaggcaggcagcctcttagag 1019

Qy   5939 gttagccggcaggaagtgggtttccgcaactttgaactaaaagacgggattatgtacctt 5998
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1020 gttagccggcaggaagtgggtttccgcaactttgaactaaaagacgggattatgtacctt 1079

Qy   5999 aacggccagcggatcgtcttcaaggggccaaccggcacgaatttgacagtaagttgggt 6058
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1080 aacggccagcggatcgtcttcaaggggccaaccggcacgaatttgacagtaagttgggt 1139

Qy   6059 cgggctatcacggaagaggatatgatctgggacatcaagaccatgaagcgaagcaacatc 6118
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 cgggctatcacggaagaggatatgatctgggacatcaagaccatgaagcgaagcaacatc 1199

Qy   6119 aatgctgtccgctgctctcactacccgaaccagtccctcttttaccggctctgtgacaag 6178
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1200 aatgctgtccgctgctctcactacccgaaccagtccctcttttaccggctctgtgacaag 1259

Qy   6179 tacggcctttacgtcattgatgaagctaacctggaaagccacggcacctgggaaaaagtg 6238
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1260 tacggcctttacgtcattgatgaagctaacctggaaagccacggcacctgggaaaaagtg 1319

Qy   6239 gggggcacgaagatcctagcttcaatgttccaggcgatgaccagcattggctgggagcc 6298
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1320 gggggcacgaagatcctagcttcaatgttccaggcgatgaccagcattggctgggagcc 1379

Qy   6299 agcttatcccgggtgaagaacatgatggctcgggacaagaaccatgcttcaatcctaatc 6358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1380 agcttatcccgggtgaagaacatgatggctcgggacaagaaccatgcttcaatcctaatc 1439

Qy   6359 tggtctttaggcaatgagtcttacgccggcactgtctttgcccaaatggctgattacgtc 6418
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1440 tggtctttaggcaatgagtcttacgccggcactgtctttgcccaaatggttgattacgtc 1499

Qy   6419 cggaaggctgatccgacccgggttcagcactatgaaggggtgacccacaaccggaagttt 6478
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 cggaaggctgatccgacccgggttcagcactatgaaggggtgacccacaaccggaagttt 1559

Qy   6479 gacgacgccacccagattgaaagccggatgtatgctccggccaaggtaattgaagaatac 6538
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1560 gacgacgccacccagattgaaagccggatgtatgctccggccaaggtaattgaagaatac 1619

Qy   6539 ttgaccaataaaccagccaagccatttatctcagttgaatacgctcacgccatgggcaac 6598
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   1620 ttgaccaataaaccagccaagccatttatctcagttgaatac-ctcacgccatgggcaac 1678

Qy   6599 tccgtcggtgacctggccgcctacacggccctggaaaaatacccccactaccagggcggc 6658
           |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db   1679 tccgtcggtgacctg--cgcctacacggccctggaaaaatacccccactaccagggcggc 1736

Qy   6659 ttcatctgggactggattgaccaaggactggaaaaagacgggcacctgctttatgggggc 6718
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1737 ttcatctgggactggattgaccaaggactggaaaaagacgggcacctgctttatgggggc 1796
```

```
Qy   6719 gacttcgatgaccggccaaccgactatgaattctgcgggaacggcctggtctttgctgac 6778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1797 gacttcgatgaccggccaaccgactatgaattctgcgggaacggcctggtctttgctgac 1856

Qy   6779 cggactgaatcgccgaaactggctaatgtcaaggccctttacgccaaccttaagttagaa 6838
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1857 cggactgaatcgccgaaactggctaatgtcaaggccctttacgccaaccttaagttagaa 1916

Qy   6839 gtaaaagatgggcagctcttcctcaaaaacgacaatttatttaccaacagctcatcttac 6898
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1917 gtaaaagatgggcagctcttcctcaaaaacgacaatttatttaccaacagctctacttac 1976

Qy   6899 tacttcttgactagtcttttggtcgatggcaagttgacctaccagagccggcctctgacc 6958
          ||||||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db   1977 tacttcttgactagtcaaaaggtcgatggcaagttgacctaccagagccggcctctgacc 2036

Qy   6959 tttggcctggagcctggcgaatccgggacctttgccctgccttggccggaagtcgctgat 7018
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2037 tttggcctggagcctggcgaatccgggacctttgccctgccttggccggaagtcgctgat 2096

Qy   7019 gaaaaaggggaggtcgtctaccgggtaacggcccacttaaaagaagacttgccttgggcg 7078
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2097 gaaaaaggagaggtcgtctaccgggtaacggcccacttaaaagaagacttgccttgggcg 2156

Qy   7079 gatgagggcttcactgtggctgaagcagaagaagtagctcaaaagctgccggaatttaag 7138
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2157 gatgagggcttcactgtggctgaagcagaagaagtagctcaaaagctgccggaatttaag 2216

Qy   7139 ccggaagggcggccagatttagttgattccgactacaacctaggcctgaaaggaaataac 7198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2217 ccggaagggcggccagatttagttgattccgactacaacctaggcctgaaaggaaataac 2276

Qy   7199 ttccaaattctcttctccaaggtcaagggctggccggtttccctcaagtatgccggtagg 7258
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2277 ttccaaattctcttctccaaggtcaagggctggccggtttccctcaagtatgccggtagg 2336

Qy   7259 gaatacttgaagcggctgccggaatttaccttctggcgggccctgacggacaacgaccgg 7318
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2337 gaatacttgaagcggctgccggaatttaccttctggcgggccctgacggacaacgaccgg 2396

Qy   7319 ggagctggttacggctatgatctggcccggtgggaaaatgccggcaagtatgcccgcttg 7378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2397 ggagctggttacggctatgatctggcccggtgggaaaatgccggcaagtatgcccgcttg 2456

Qy   7379 aaagacatcagctgcgaggtcaaggaagactccgttttggtcaagactgcctttacgttg 7438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2457 aaagacatcagctgcgaggtcaaggaagactccgttttggtcaagactgcctttacgttg 2516

Qy   7439 cctgtcgccttaaagggtgatttaaccgtgacctatgaagtcgatggacggggcaagatt 7498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2517 cctgtcgccttaaagggtgatttaaccgtgacctatgaagtcgatggacggggcaagatt 2576

Qy   7499 gctgtaacagctgacttcccaggcgcggaagaagctggtctcttgccagcctttggcttg 7558
          ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   2577 gctgtaacagctgacttcccaggcgcggaagaagccggtctcttgccagcctttggcttg 2636

Qy   7559 aacctggccctgccaaaagaactgaccgattaccgctactatggtctgggacctaatgag 7618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2637 aacctggccctgccaaaagaactgaccgattaccgctactatggtctgggacctaatgag 2696

Qy   7619 agctacccagaccgcttggaaggtaattacctgggcatctaccagggagcggtaaaaaag 7678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2697 agctacccagaccgcttggaaggtaattacctgggcatctaccagggagcggtaaaaaag 2756

Qy   7679 aactttagcccata---tcgtccgcaggaaacgggcaaccggagcaaggttcgctggtac 7735
          ||||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db   2757 aactttagcccatacctgcgtccgcaggaaacgggcaaccggagcaaggttcgctggtac 2816

Qy   7736 cagctctttgatgaaaagggcggcttggaatttacggccaatggggcagacttgaacttg 7795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2817 cagctctttgatgaaaagggcggcttggaatttacggccaatggggcagacttgaacttg 2876
```

```
Qy   7796 tctgctttgccatattctgccgcccaaattgaagcagcggaccacgcttttgaactgact 7855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2877 tctgctttgccatattctgccgcccaaattgaagcagcggaccacgcttttgaactgact 2936

Qy   7856 aacaattacacttgggttagagccttaagcgcccagatgggggtcggcggggatgactcc 7915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2937 aacaattacacttgggttagagccttaagcgcccagatgggggtcggcggggatgactcc 2996

Qy   7916 tgggggcagaaggtccacccggaattctgcctggatgctcaaaaagcccgccagcttcgc 7975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||  | |||
Db   2997 tgggggcagaaggtccacccggaattctgcctggatgctcaaaaagcccgccacgtccgc 3056

Qy   7976 ctggtgattcagccccttttactaaaataaatgctacaattgacttaacaggatgaaatt 8035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3057 ctggtgattcagccccttttactaaaataaatgctacaattgacttaacaggatgaaatt 3116

Qy   8036 ttagtaaaagcaaagcgagtgaggaagatggcaacgatcagagaagtgccaaggcagccg 8095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3117 ttagtaaaagcaaagcgagtgaggaagatggcaacgatcagagaagtgccaaggcagccg 3176

Qy   8096 gcgtgtcgctagcgacggt 8114
          ||||||||||||||||||
Db   3177 gcgtgtcgctagcgacggt 3195
```

```
RESULT   3
AAV40006
ID   AAV40006 standard; DNA; 3853 BP.
XX
AC   AAV40006;
XX
DT   15-FEB-1999  (first entry)
XX
DE   Plasmid pCTM.
XX
KW   E2F; transcription factor; human; retinoblastoma protein RB;
KW   bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW   thyroid hyperplasia; Grave's disease; psoriasis;
KW   benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW   peripheral vascular disease; therapy; plasmid pCTM; ss.
XX
OS   Chimeric - cytomegalovirus.
OS   Chimeric - mastadenovirus.
OS   Chimeric - bacteriophage T7.
OS   Chimeric - bacteriophage SP6.
OS   Chimeric - rhesus macaque polyoma virus.
OS   Chimeric - Bos taurus.
XX
FH   Key             Location/Qualifiers
FT   promoter        209..864
FT                   /*tag= a
FT                   /note= "CMV promoter"
FT   misc_feature    907..1131
FT                   /*tag= b
FT                   /function= tripartite leader sequence
FT   promoter        1132..1149
FT                   /*tag= c
FT                   /note= "SP6 promoter"
FT   misc_feature    1679..3853
FT                   /*tag= d
FT                   /note= "pUC19 backbone H3 to AatII"
FT   CDS             complement (2857..3717)
FT                   /*tag= e
FT                   /note= "AMP-ORF"
XX
PN   WO9821228-A1.
XX
PD   22-MAY-1998.
XX
PF   13-NOV-1997;   97WO-US21821.
XX
PR   14-FEB-1997;   97US-0801092.
PR   15-NOV-1996;   96US-0751517.
```

XX
PA (CANJ-) CANJI INC.
XX
PI Antelman D, Gregory RJ, Wills KN;
XX
DR WPI; 1998-297858/26.
XX
PT New fusion polypeptide of, e.g. transcription factor - used to
PT treat, e.g. hyper-proliferative disease such as cancer and
PT restenosis
XX
PS Example 1; Fig 4; 91pp; English.
XX
CC This is the nucleotide sequence of pCTM, a plasmid which contains
CC a CMV promoter, a tripartite adenovirus leader flanked by T7 and
CC SP6 promoters, and a multiple cloning site with a bovine growth
CC hormone polyA site and downstream SV40 polyA site. It has been
CC used as a vector for the expression of fusion proteins of the
CC invention that comprise retinoblastoma protein (BP, see AAW62465) and
CC E2F transcription factor (see AAW62464). Such fusion proteins,
CC particularly expressed from gene therapy vectors, are used to treat
CC hyperproliferative conditions, specifically cancer (particularly of
CC the bladder) or restenosis. They are more effective in repressing
CC transcription of the E2F promoter than RB alone and cause
CC cell-cycle arrest in a variety of cells.
XX
SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 other;

Query Match 15.3%; Score 1245.2; DB 19; Length 3853;
Best Local Similarity 74.2%; Pred. No. 1.3e-177;
Matches 1914; Conservative 0; Mismatches 93; Indels 573; Gaps 4;

```
Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 269

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 329

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 389

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 449

Qy    241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 509

Qy    301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 569

Qy    361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 629

Qy    421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 689

Qy    481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 749

Qy    541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 809

Qy    601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 869

Qy    661 actcactatagggagacccaagctt---------------------------------- 685
          |||||||||||||||||||||||||
Db    870 actcactatagggagacccaagcttcgcgcgggtaccactctcttccgcatcgctgtctg 929

Qy    686 ---------------------------------------------------------- 685

Db    930 cgagggccagctgttgggctcgcggttgaggacaaactcttcgcggtctttccagtactc 989

Qy    686 -------------------ggtaccgagctcggatccactagtaacggccgccagtgtgc 726
                              |   |||| |     ||| | |   | || |   || | |
Db    990 ttggatcggaaacccgtcggcctccgaacggtactccgccaccgagggacctgagcgagt 1049

Qy    727 tggaattctgcagatatccatcacactggcggccgc-------------------tcgag 767
            | ||    | |||    |   | || | ||| ||                   ||| |
Db   1050 ccgcatcgaccggatcggaaaacctctcgaggcggccgctgcagtctagacgaattcgcg 1109

Qy    768 catgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctgatcagc 827
            | |         ||||||||||||||||||||||||||||||||||||||||||||||
Db   1110 tacgatatcgatgggccctattctatagtgtcacctaaatgctagagctcgctgatcagc 1169

Qy    828 ctcgactgtgccttctagttgccagccatctgttgtttgcccctccccccgtgccttcctt 887
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1170 ctcgactgtgccttctagttgccagccatctgttgtttgcccctccccccgtgccttcctt 1229

Qy    888 gaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgca 947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1230 gaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgca 1289

Qy    948 ttgtctgagtaggtgtcattctattctgggggggtgggggtgggggcaggacagcaagggga 1007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1290 ttgtctgagtaggtgtcattctattctgggggggtgggggtgggggcaggacagcaagggga 1349

Qy   1008 ggattgggaagacaatagcaggcatgctgggga---------------------------- 1040
          ||||||||||||||||||  |  |||   |
Db   1350 ggattgggaagacaatagccgaaatgaccgaccaagcgacgcccaacctgccatcacgag 1409

Qy   1041 ---------------------------------------------------------- 1040

Db   1410 atttcgattccaccgccgccttctatgaaaggttgggcttcggaatcgttttccgggacg 1469

Qy   1041 ---------------------------------------------------------- 1040

Db   1470 ccggctggatgatcctccagcgcggggatctcatgctggagttcttcgcccaccccaact 1529

Qy   1041 ---------------------------------------------------------- 1040

Db   1530 tgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaata 1589

Qy   1041 ---------------------------------------------------------- 1040

Db   1590 aagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatc 1649

Qy   1041 ---------------------------------------------------------- 1040

Db   1650 atgtctgtataccgtcgacctctagctagagcttggcgtaatcatggtcatagctgtttc 1709

Qy   1041 ---------------------------------------------------------- 1040

Db   1710 ctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagt 1769

Qy   1041 ----------------------------------------------tgcggtgggctctatgg 1057
                                                       |||| || ||||  ||
Db   1770 gtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgc 1829

Qy   1058 cttct----------gaggcggaaagaaccagctgcattaatgaatcggccaacgcgcgg 1107
          |  ||          ||  |     |  ||||||||||||||||||||||||||||||
Db   1830 ccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcgg 1889

Qy   1108 ggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgct 1167
```

```
Db    1890 ggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgct 1949

Qy    1168 cggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatcca 1227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1950 cggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatcca 2009

Qy    1228 cagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccagga 1287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2010 cagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccagga 2069

Qy    1288 accgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatc 1347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2070 accgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatc 2129

Qy    1348 acaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg 1407
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2130 acaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg 2189

Qy    1408 cgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggat 1467
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2190 cgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggat 2249

Qy    1468 acctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggt 1527
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2250 acctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggt 2309

Qy    1528 atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttc 1587
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2310 atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttc 2369

Qy    1588 agcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacg 1647
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2370 agcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacg 2429

Qy    1648 acttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg 1707
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2430 acttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg 2489

Qy    1708 gtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttg 1767
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2490 gtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttg 2549

Qy    1768 gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccg 1827
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2550 gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccg 2609

Qy    1828 gcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgca 1887
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2610 gcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgca 2669

Qy    1888 gaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtgga 1947
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2670 gaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtgga 2729

Qy    1948 acgaaaactcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
           |||||||||||||||||||||||||||||||||  ||  | |      |  |  ||||
Db    2730 acgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctaga 2789


RESULT   4
AAA53878
ID   AAA53878 standard; DNA; 5306 BP.
XX
AC   AAA53878;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Expression vector pRIG-22.
XX
KW   Vector; endogenous gene; activation; over-expression;
KW   erythropoietin; growth hormone; drug discovery;
KW   granulocyte colony stimulating factor; ds.
XX
OS   Synthetic.
XX
PN   WO200049162-A2.
XX
PD   24-AUG-2000.
XX
PF   22-FEB-2000; 2000WO-US04429.
XX
PR   19-FEB-1999;   99US-0253022.
PR   08-MAR-1999;   99US-0263814.
PR   26-MAR-1999;   99US-0276820.
XX
PA   (ATHE-) ATHERSYS INC.
XX
PI   Harrington JJ,  Sherf B,  Rundlett S;
XX
DR   WPI; 2000-549276/50.
XX
PT   Non-targeted activation of endogenous genes, e.g. for the production of
PT   erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT   proteins and for drug discovery
XX
PS   Disclosure; Fig 35; 240pp; English.
XX
CC   New methods, vectors and cells are described for non-targeted
CC   activation and over-expression of endogenous genes. The vector
CC   constructs comprise transcriptional regulatory sequences (TRS) and
CC   unpaired splice donor sequences (USDS), preferably the vectors
CC   comprise (in sequential order) a TRS, an USDS, a rare cutting
CC   restriction site (RCRS) and a linearization site (LS) with a second
CC   TRS linked to a selectable marker (SM) lacking a polyadenylation
CC   signal. The methods, vectors and cells comprising the vectors may
CC   be used for the non-targeted activation and over-expression of
CC   endogenous genes, e.g. for the production of proteins (including
CC   erythropoietin, growth hormone or granulocyte-colony stimulating
CC   factor) and drug discovery. The advantage of these methods are that
CC   endogenous genes including those associated with human disease and
CC   development, may be activated and isolated without prior knowledge
CC   of the sequence structure, function or expression profile of
CC   the genes being known.
XX
SQ   Sequence 5306 BP; 1303 A; 1365 C; 1383 G; 1255 T; 0 other;


  Query Match            14.2%;  Score 1152.4;  DB 21;  Length 5306;
  Best Local Similarity  76.4%;  Pred. No. 8.2e-164;
  Matches 1524;  Conservative    0;  Mismatches  396;  Indels   74;  Gaps    6;

Qy      23 gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcata 82
           ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 gttggcattgattattgactagttattaatagtaatcaattacggggtcattagttcata 199

Qy      83 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 259

Qy     143 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 319

Qy     203 ggactttccattgacgtcaatgggtggactatttacggtaaactgcccacttggcagtac 262
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     320 ggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtac 379

Qy     263 atcaagtgtatcatatgccaagtacgcccccattgacgtcaatgacggtaaatggcccg 322
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     380 atcaagtgtatcatatgccaagtccgcccccattgacgtcaatgacggtaaatggcccg 439

Qy     323 cctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacg 382
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     440 cctggcattatgcccagtacatgaccttacgggactttcctacttggcagtacatctacg 499
```

```
Qy    383 tattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggat 442
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    500 tattagtcatcgctattaccatggtgatgcggttttggcagtacaccaatgggcgtggat 559

Qy    443 agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgt 502
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560 agcgctttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgt 619

Qy    503 tttggcaccaaaatcaacgggactttccaaaatgtcgtaacaact--------ccgcccc 554
          ||||||||||||||||||||||||||||||||||||||||||||        |||||||
Db    620 tttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactgcgatcgcccgcccc 679

Qy    555 attgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctctg 614
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db    680 gttgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgttt 739

Qy    615 gctaactagagaacccactgcttactggcttatcgaaattaatacgactcactatagggga 674
            | |   | |   |||||  | ||| |          | |   | |  |
Db    740 agtgaaccgtcagatcactgaattctgacgacctactgattaacggccatagaggcctcc 799

Qy    675 gacccaagcttggtaccgagctcggatccactagtaacggccgccagtgtgctg-gaatt 733
            |  ||   |  ||  ||   |       |    ||      |     | ||  ||||
Db    800 tgcagaactgtcttagtgacaactatcgatttccacacattatacgagccgatgttaatt 859

Qy    734 ctgcagatatccatcacactggcggccgctcgagcatgcatctagagggccctattctat 793
            |  | |  ||   ||     | |    |  |   || ||  |   | | | | |
Db    860 gtcaacagctcatgcatgacgtcccgggagcagacaagccccaccatggctcgagtaata 919

Qy    794 agtgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagc 853
            |  ||||   |  || | || |   |  ||     |    | |
Db    920 cgactcactatagggcgacaggtgagtactcgctaccttaagagag-------------- 965

Qy    854 catctgttgtttgcccctccccgtgccttccttgaccctggaaggtgccactcccactg 913
                                  ||||  || | |  |     ||   |     |
Db    966 ------------------------gcctatctggccagttagcagtcgaagaaagaagt 1000

Qy    914 tcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattc 973
          |        |||     |    ||   |  |      |||   |       | ||
Db   1001 ttaagagagccgaaacaagcgctcatgagcccgaagtggcgagcccgatcttccccatcg 1060

Qy    974 tgggggggtggggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatg 1033
            |  |  || |        | ||||||  |      | | |  |   ||  | | || |
Db   1061 gtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacg 1120

Qy   1034 ctggggatgcggtgggctctatggcttctgaggcggaaagaaccagctgcattaatgaat 1093
            || |           |  ||| |    |    | || ||  |    |
Db   1121 atgcg-------------tccggcgtagaggatccacaggacgggtgtggtcgccatga 1166

Qy   1094 cggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcac 1153
            ||  |  ||   |  |   |  | ||| |  | ||     |  ||  |  |     |
Db   1167 tcgcgtagtcgatagtggctccaagtagcgaagcgagcaggactgggcggcggccaaagc 1226

Qy   1154 tgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggt 1213
            | |   | | |||| |  |  ||| ||||   ||   |  || |       |  ||| |
Db   1227 ggtcggacagtgctccg-agaacgggtgcgcatagaaattgcatcaacgcatatagcgct 1285

Qy   1214 aatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggcca 1273
          |   | |  |  |     |    |  | |           ||||||||||||||||||
Db   1286 agatccttgctagagtcgagatctgtcgagc----------catgtgagcaaaaggcca 1334

Qy   1274 gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccc 1333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1335 gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccc 1394

Qy   1334 ccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact 1393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1395 ccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact 1454

Qy   1394 ataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccct 1453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1455 ataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccct 1514

Qy   1454 gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatg 1513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db   1515 gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatag 1574

Qy   1514 ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgca 1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1575 ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgca 1634

Qy   1574 cgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaa 1633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1635 cgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaa 1694

Qy   1634 cccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagc 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1695 cccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagc 1754

Qy   1694 gaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactag 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1755 gaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactag 1814

Qy   1754 aaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttgg 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1815 aaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttgg 1874

Qy   1814 tagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagca 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1875 tagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagca 1934

Qy   1874 gcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtc 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1935 gcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtc 1994

Qy   1934 tgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacatatt 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||   ||  | |
Db   1995 tgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaag 2054

Qy   1994 tgaatgtatttaga 2007
                |  |  ||||
Db   2055 gatcttcacctaga 2068
```

RESULT 5
AAV32374/c
ID AAV32374 standard; DNA; 14455 BP.
XX
AC AAV32374;
XX
DT 29-SEP-1998 (first entry)
XX
DE Complete sequence of the pE1/Fiber plasmid.
XX
KW Circular; adenovirus type 5; pE1/Fiber plasmid; structural protein;
KW complementation; fiber protein; gene therapy; HIV; tumour; AD5;
KW early gene; Huntington's disease; Tay-Sachs disease;
KW sickle cell disease; E1 regulatory protein; ds.
XX
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	misc_feature	1460..4998
FT		/*tag= a
FT		/note= "AD5 E1 regulatory gene"
FT	misc_feature	complement (10922..14223)
FT		/*tag= b
FT		/note= "AD5 fiber gene consisting of a CMV promoter at 5' end of this gene"

XX
PN WO9813499-A2.
XX
PD 02-APR-1998.
XX
PF 24-SEP-1997; 97WO-EP05251.
XX
PR 25-SEP-1996; 96US-0719806.

```
XX
PA   (NOVS ) NOVARTIS AG.
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Memerow GR,  Von Seggern DJ;
XX
DR   WPI; 1998-230709/20.
XX
PT   Adenoviral vectors - which lack DNA encoding for structural protein
PT   or fibre protein used particularly for gene therapy
XX
PS   Example 1; Pages 112-131; 170pp; English.
XX
CC   The present sequence is that of a pE1/Fiber plasmid used in the
CC   method of the invention.  The plasmid contains an adenovirus type 5 (AD5)
CC   fiber gene controlled by a CMV promoter, an AD5 E1 gene and a pMAM
CC   backbone.  The invention provides adenoviral vectors having deletions of
CC   all or part of various gene sequences encoding adenoviral structural
CC   proteins and/or early region proteins.  Deletions in these proteins would
CC   allow a reduced risk of wild-type virus contamination and would also
CC   allow packaging of foreign DNA in such vectors for a variety of
CC   diagnostic and therapeutic applications.  The adenoviral vectors having
CC   deletions in the structural and/or early gene regions are produced by
CC   cellular complementation of these adenoviral genes.  Therefore, the
CC   pE1/Fiber plasmid was used as a complementation plasmid which was
CC   introduced into a host cell line where parts of the fiber and E1 gene
CC   region would be stably inserted into the host cell chromosomes.  The
CC   resulting E1/fiber gene deficient plasmid can be used as a gene delivery
CC   vector.  The vectors can be used for diagnosis or gene therapy, e.g. for
CC   treating conditions characterised by hyper-proliferative cells (e.g.
CC   tumours), genetic diseases (e.g. Huntington's disease, Tay-Sachs disease,
CC   or sickle cell disease), or infections (e.g. HIV infection).  They can
CC   also be used for in vitro production of biologically active proteins.
XX
SQ   Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;
```

```
  Query Match             14.1%;  Score 1148.2;  DB 19;  Length 14455;
  Best Local Similarity   95.5%;  Pred. No. 3e-163;
  Matches 1196;  Conservative    0;  Mismatches   48;  Indels    9;  Gaps    1;

Qy     755 gcggccgctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagag 814
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11236 GCGGCCGCTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAG 11177

Qy     815 ctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccc 874
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11176 CTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC 11117

Qy     875 ccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgagg 934
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11116 CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGG 11057

Qy     935 aaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcagg 994
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11056 AAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGG 10997

Qy     995 acagcaaggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctcta 1054
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10996 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 10937

Qy    1055 tggcttctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagagg 1114
           |||||||||||||| |          ||| |    |||    ||  |    |  ||  |
Db   10936 TGGCTTCTGAGGCGGTA---------TGCGGTGTGAAATACCGCACAGATGCGTAAGGAG 10886

Qy    1115 cggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgt 1174
             |  || ||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db   10885 AAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT 10826

Qy    1175 tcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatc 1234
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10825 TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC 10766
```

```
Qy    1235 aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaa 1294
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10765 AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAA 10706

Qy    1295 aaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaa 1354
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10705 AAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA 10646

Qy    1355 tcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcc 1414
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10645 TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCC 10586

Qy    1415 ccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtc 1474
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10585 CCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC 10526

Qy    1475 cgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcag 1534
           ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   10525 CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAG 10466

Qy    1535 ttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccccgttcagcccga 1594
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10465 TTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA 10406

Qy    1595 ccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatc 1654
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10405 CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC 10346

Qy    1655 gccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctac 1714
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10345 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC 10286

Qy    1715 agagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctg 1774
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10285 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTG 10226

Qy    1775 cgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca 1834
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10225 CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA 10166

Qy    1835 aaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaa 1894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10165 AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA 10106

Qy    1895 aggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa 1954
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10105 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA 10046

Qy    1955 ctcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
           ||||||||||||||||||||||||||   ||  | |      |  |  ||||
Db   10045 CTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 9993
```

```
RESULT   6
AAA59050/c
ID   AAA59050 standard; DNA; 14455 BP.
XX
AC   AAA59050;
XX
DT   07-NOV-2000  (first entry)
XX
DE   Nucleotide sequence of the E1/fiber-expressing plasmid pE1/Fiber.
XX
KW   Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
KW   E1 gene; fiber gene; ss.
XX
OS   Synthetic.
OS   Adenovirus type 5.
XX
PN   WO200042208-A1.
XX
PD   20-JUL-2000.
XX
```

```
PF   14-JAN-2000; 2000WO-EP00265.
XX
PR   14-JAN-1999;   99US-0115920.
XX
PA   (NOVS ) NOVARTIS AG.
PA   (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Nemerow GR,  Von Seggern DJ,  Hallenbeck PL,  Stevenson SC;
PI   Skripchenko Y;
XX
DR   WPI; 2000-476068/41.
XX
PT   New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT   for producing high-capacity and targeted vectors for adenovirus-based
PT   gene therapy -
XX
PS   Example 1; Page 160-164; 212pp; English.
XX
CC   The specification describes a nucleic acid molecule comprising an
CC   adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC   comprising two different TPL exons or three same or different TPL
CC   exons. The nucleic acid is used to produce an adenovirus vector particle,
CC   deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC   vectors, target an adenovirus vector to a cell, produce a modified
CC   adenovirus, deliver a heterologous gene to an animal and produce a
CC   gutless adenoviral vector particle. The present sequence represents
CC   a E1/fiber-expressing plasmid, which was used for complementation of
CC   E1/fiber-gene deleted adenoviruses.
XX
SQ   Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;
```

Query Match 14.1%; Score 1148.2; DB 21; Length 14455;
Best Local Similarity 95.5%; Pred. No. 3e-163;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;

```
Qy    755 gcggccgctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagag 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11236 GCGGCCGCTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAG 11177

Qy    815 ctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccc 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11176 CTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCC 11117

Qy    875 ccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgagg 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11116 CCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGG 11057

Qy    935 aaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcagg 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11056 AAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGG 10997

Qy    995 acagcaaggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctcta 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10996 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 10937

Qy   1055 tggcttctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagagg 1114
          |||||||||||||| |         ||| |    ||| |    || |    | || |
Db  10936 TGGCTTCTGAGGCGGTA---------TGCGGTGTGAAATACCGCACAGATGCGTAAGGAG 10886

Qy   1115 cggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgt 1174
              |  || ||  ||||||||||||||||||||||||||||||||||||||||||||||
Db  10885 AAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT 10826

Qy   1175 tcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatc 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10825 TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC 10766

Qy   1235 aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaa 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10765 AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAA 10706

Qy   1295 aaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaa 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10705 AAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA 10646

Qy   1355 tcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcc 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10645 TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCC 10586

Qy   1415 ccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtc 1474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10585 CCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC 10526

Qy   1475 cgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcag 1534
          ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db  10525 CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAG 10466

Qy   1535 ttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccga 1594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10465 TTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA 10406

Qy   1595 ccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatc 1654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10405 CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC 10346

Qy   1655 gccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctac 1714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10345 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC 10286

Qy   1715 agagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctg 1774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10285 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTG 10226

Qy   1775 cgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca 1834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10225 CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA 10166

Qy   1835 aaccaccgctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaa 1894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10165 AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA 10106

Qy   1895 aggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa 1954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10105 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA 10046

Qy   1955 ctcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          |||||||||||||||||||||||||||  ||  | |      | |  ||||
Db  10045 CTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 9993
```

```
RESULT   7
AAA53872
ID   AAA53872 standard; DNA; 5377 BP.
XX
AC   AAA53872;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Expression vector pRIG14.
XX
KW   Vector; endogenous gene; activation; over-expression;
KW   erythropoietin; growth hormone; drug discovery;
KW   granulocyte colony stimulating factor; ds.
XX
OS   Synthetic.
XX
PN   WO200049162-A2.
XX
PD   24-AUG-2000.
XX
PF   22-FEB-2000; 2000WO-US04429.
XX
PR   19-FEB-1999;   99US-0253022.
PR   08-MAR-1999;   99US-0263814.
```

PR 26-MAR-1999; 99US-0276820.
XX
PA (ATHE-) ATHERSYS INC.
XX
PI Harrington JJ, Sherf B, Rundlett S;
XX
DR WPI; 2000-549276/50.
XX
PT Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery
XX
PS Disclosure; Fig 29; 240pp; English.
XX
CC New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDS), preferably the vectors
CC comprise (in sequential order) a TRS, an USDS, a rare cutting
CC restriction site (RCRS) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
CC the genes being known.
XX
SQ Sequence 5377 BP; 1248 A; 1465 C; 1437 G; 1225 T; 2 other;

Query Match 14.0%; Score 1140; DB 21; Length 5377;
Best Local Similarity 74.3%; Pred. No. 5.7e-162;
Matches 1684; Conservative 0; Mismatches 300; Indels 284; Gaps 8;

```
Qy     23 gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcata 82
Db   2058 gttggcattgattattgactagttattaatagtaatcaattacggggtcattagttcata 2117

Qy     83 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 142
Db   2118 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 2177

Qy    143 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 202
Db   2178 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 2237

Qy    203 ggactttccattgacgtcaatgggtggactatttacggtaaactgcccacttggcagtac 262
Db   2238 ggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtac 2297

Qy    263 atcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccg 322
Db   2298 atcaagtgtatcatatgccaagtccgccccctattgacgtcaatgacggtaaatggcccg 2357

Qy    323 cctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacg 382
Db   2358 cctggcattatgcccagtacatgaccttacgggactttcctacttggcagtacatctacg 2417

Qy    383 tattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggat 442
Db   2418 tattagtcatcgctattaccatggtgatgcggttttggcagtacaccaatgggcgtggat 2477

Qy    443 agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgt 502
Db   2478 agcgg-ttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgt 2536

Qy    503 tttggcaccaaaatcaacgggactttccaaaatgtcgtaacaact--------ccgcccc 554
Db   2537 tttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactgcgatcgcccgcccc 2596

Qy    555 attgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctctg 614
Db   2597 gttgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgttt 2656

Qy    615 gctaactagagaacccact----------------------gcttactggcttatc 648
Db   2657 agtgaaccgtcagatcactagaagtttattgcggtagtttatcacagttaaattgctaac 2716

Qy    649 gaaattaatacgactcactatagggagacccaagcttggtaccgagctcggatccactag 708
Db   2717 gcagtcagtgcttctgacacaacagtctcgaacttaagctgcagtgactctcttaattaa 2776

Qy    709 taacggccgccagtgtg---------------------------------------- 725
Db   2777 ccaccgctacaggtgagtactcggatctgctaccttaagagaggcctatctggccagtta 2836

Qy    726 ---------------------------------------------------------- 725
Db   2837 gcagtcgaagaaagaagtttaagagagccgaaacaagcgctcatgagcccgaagtggcga 2896

Qy    726 ------------------------ctggaattctgcagatatccatcacactggcggcc 760
Db   2897 gcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcg 2956

Qy    761 gctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagagctcgc- 819
Db   2957 ccggtgatgccggccacgatgcgtccggcgtagaggatccacaggacgggtgtggtcgcc 3016

Qy    820 ---tgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccc 876
Db   3017 atgatcgcgtagtcgatagtggctccaagtagcgaagcgagcaggactgggcggcggcca 3076

Qy    877 gtgccttccttga---------------------------------------------- 889
Db   3077 aagcggtcggacagtgctccgagaacgggtgcgcatagaaattgcatcaacgcatatagc 3136

Qy    890 ------------------------ccctggaaggtgccactcccactgtcctttccta 923
Db   3137 gctagatccttgctagagtcgaggccgccaccgcggtggagctccagctttgttccctt 3196

Qy    924 ataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggg----- 978
Db   3197 tagtgagggttaatttcgagcttggcgtaatcatggtcatagctgtttcctgtggtgaaa 3256

Qy    979 ----------------------------ggtggggtggggcaggacagcaaggggagg 1009
Db   3257 ttgttatccgctcacaattccacaacatacgagccggaagcataaagtgtaaagcctggg 3316

Qy   1010 attgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttct------- 1062
Db   3317 gtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagt 3376

Qy   1063 ---gaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119
Db   3377 cgggaaacctgtcgtgccagctgcattaatgaatccgccaacgcgcggggagaggcggtt 3436

Qy   1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179
Db   3437 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 3496

Qy   1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239
Db   3497 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 3556

Qy   1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299
Db   3557 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 3616

Qy   1300 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 1359
Db   3617 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 3676
```

```
Qy   1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3677 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 3736

Qy   1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3737 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 3796

Qy   1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539
          |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   3797 ttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgg 3856

Qy   1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgct 1599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3857 tgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgct 3916

Qy   1600 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3917 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 3976

Qy   1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3977 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 4036

Qy   1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4037 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 4096

Qy   1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4097 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 4156

Qy   1840 ccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4157 ccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 4216

Qy   1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4217 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 4276

Qy   1960 gttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          |||||||||||||||||||  ||  | |      |  |  ||||
Db   4277 gttaagggattttggtcatgagattatcaaaaaggatcttcacctaga 4324
```

RESULT 8
AAA53870
ID AAA53870 standard; DNA; 4645 BP.
XX
AC AAA53870;
XX
DT 03-JAN-2001 (first entry)
XX
DE Expression vector pRIG21b.
XX
KW Vector; endogenous gene; activation; over-expression;
KW erythropoietin; growth hormone; drug discovery;
KW granulocyte colony stimulating factor; ds.
XX
OS Synthetic.
XX
PN WO200049162-A2.
XX
PD 24-AUG-2000.
XX
PF 22-FEB-2000; 2000WO-US04429.
XX
PR 19-FEB-1999; 99US-0253022.
PR 08-MAR-1999; 99US-0263814.
PR 26-MAR-1999; 99US-0276820.
XX
PA (ATHE-) ATHERSYS INC.
XX
PI Harrington JJ, Sherf B, Rundlett S;
XX
DR WPI; 2000-549276/50.
XX
PT Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery
XX
PS Disclosure; Fig 15; 240pp; English.
XX
CC New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDS), preferably the vectors
CC comprise (in sequential order) a TRS, an USDS, a rare cutting
CC restriction site (RCRS) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
CC the genes being known.
XX
SQ Sequence 4645 BP; 1104 A; 1247 C; 1221 G; 1072 T; 1 other;

Query Match 13.9%; Score 1131.4; DB 21; Length 4645;
Best Local Similarity 76.0%; Pred. No. 1.1e-160;
Matches 1514; Conservative 0; Mismatches 391; Indels 88; Gaps 6;

```
Qy     23 gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcata 82
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 gttggcattgattattgactagttattaatagtaatcaattacggggtcattagttcata 199

Qy     83 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 142
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    200 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 259

Qy    143 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    260 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 319

Qy    203 ggactttccattgacgtcaatgggtggactatttacggtaaactgcccacttggcagtac 262
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    320 ggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtac 379

Qy    263 atcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccg 322
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    380 atcaagtgtatcatatgccaagtccgccccctattgacgtcaatgacggtaaatggcccg 439

Qy    323 cctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacg 382
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    440 cctggcattatgcccagtacatgaccttacgggactttcctacttggcagtacatctacg 499

Qy    383 tattagtcatcgctattaccatggtgatgcggttttggcagtacatcaatgggcgtggat 442
          |||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Db    500 tattagtcatcg-tattaccatggtgatgcggttttggcagtacaccaatgggcgtggat 558

Qy    443 agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgt 502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    559 agcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttgt 618

Qy    503 tttggcaccaaaatcaacgggactttccaaaatgtcgtaacaact--------ccgcccc 554
          |||||||||||||||||||||||||||||||||||||||||||        |||||||
Db    619 tttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactgcgatcgcccgcccc 678

Qy    555 attgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctctg 614
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db    679 gttgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgttt 738
```

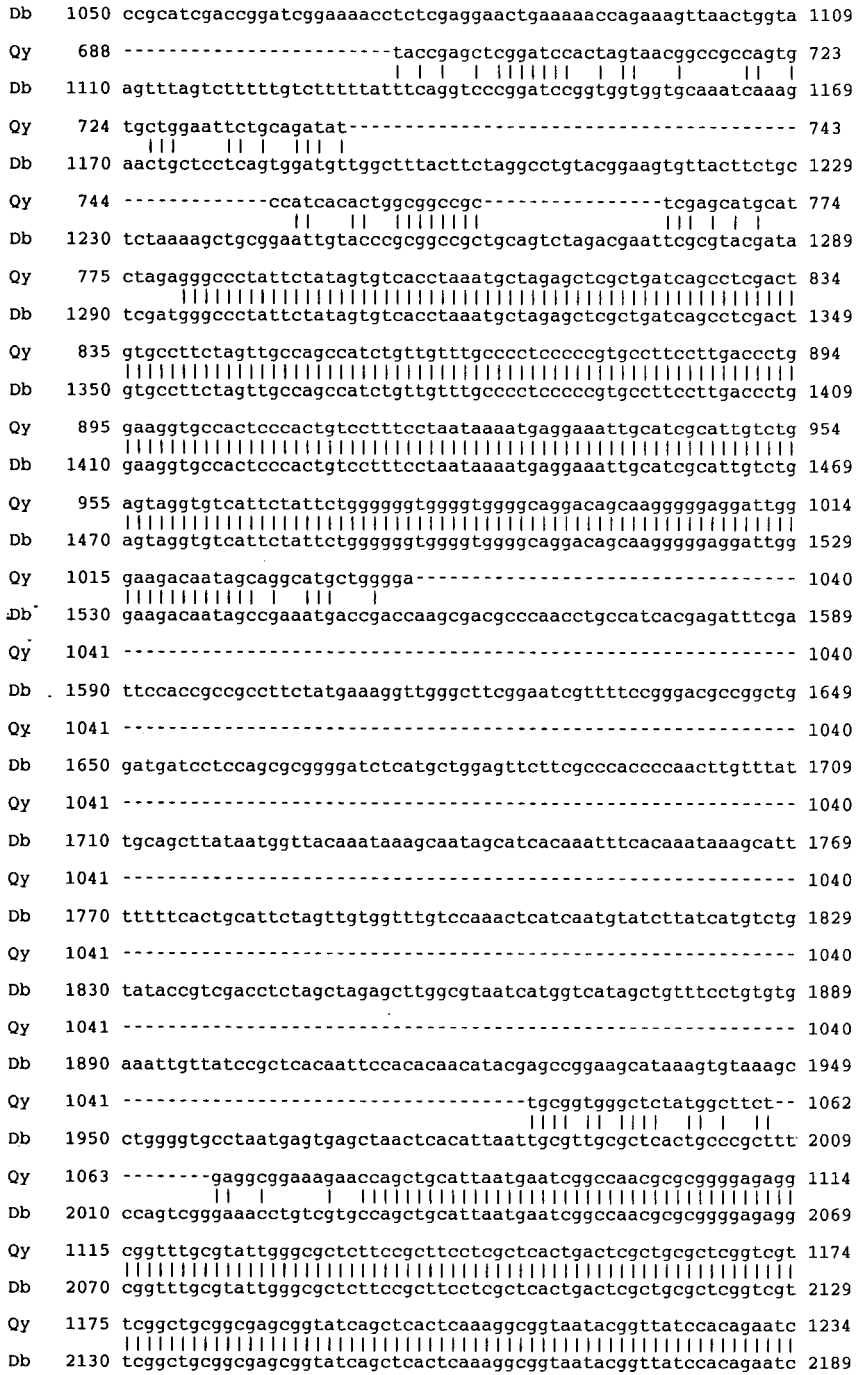
```
Qy    615 gctaactagagaacccactgcttactggcttatcgaaattaatacgactcactataggga 674
Db    739 agtgaaccgtcagatcactgaattctgac---gacctactgattaacggccatagaggcc 795

Qy    675 gacccaagcttggtaccgagctcggatccactagtaacggccgccagtgtgctggaattc 734
Db    796 tcctgcagatcactagaagctttattgcggtagtttatcacagttaaattgctaacgcag 855

Qy    735 tgcagatatccatcacactggcggccgctcgagcatgcatctagagggccctattctata 794
Db    856 tcagtgcttctgacacaacagtctcgaacttaagctgcagtgactctcttaaatccacca 915

Qy    795 gtgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagcc 854
Db    916 tggctacaggtgagtactcgctaccttaagagaggcctatctggccagttagcagtcgaa 975

Qy    855 atctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactgt 914
Db    976 gaaagaagtttaagagagccgaaacaagcgctcatgagcccgaagtggcgagcccgatct 1035

Qy    915 cctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattct 974
Db   1036 tccccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgcc------- 1088

Qy    975 gggggtggggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatgc 1034
Db   1089 ---------ggtgatgccggccacgatgcgtccggcgtagaggatccacaggacgggtgt 1139

Qy   1035 tggggatgcggtgggctctatggcttctgaggcggaaagaaccagctgcattaatgaatc 1094
Db   1140 ggtcgccatgat------------------------------cgcgtagtcgatagtgg 1168

Qy   1095 ggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcact 1154
Db   1169 ctccaactagcgaagcgagcaggactgggcggcggccaaagcggtcggacagtgctccga 1228

Qy   1155 gactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggta 1214
Db   1229 gaacgggtgcgcatagaaattgcatcaacgcatatagcgctagatccttgctagagtcga 1288

Qy   1215 atacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccag 1274
Db   1289 gatctgtcgagc----------------------------catgtgagcaaaaggccag 1319

Qy   1275 caaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccc 1334
Db   1320 caaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccc 1379

Qy   1335 cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggacta 1394
Db   1380 cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggacta 1439

Qy   1395 taaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctg 1454
Db   1440 taaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctg 1499

Qy   1455 ccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgc 1514
Db   1500 ccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagc 1559

Qy   1515 tcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcac 1574
Db   1560 tcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcac 1619

Qy   1575 gaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaac 1634
Db   1620 gaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaac 1679

Qy   1635 ccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcg 1694
Db   1680 ccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcg 1739

Qy   1695 aggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactaga 1754
Db   1740 aggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactaga 1799

Qy   1755 aggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggt 1814
Db   1800 aggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggt 1859

Qy   1815 agctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcag 1874
Db   1860 agctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaagcag 1919

Qy   1875 cagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtct 1934
Db   1920 cagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtct 1979

Qy   1935 gacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacatattt 1994
Db   1980 gacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaagg 2039

Qy   1995 gaatgtatttaga 2007
Db   2040 atcttcacctaga 2052
```

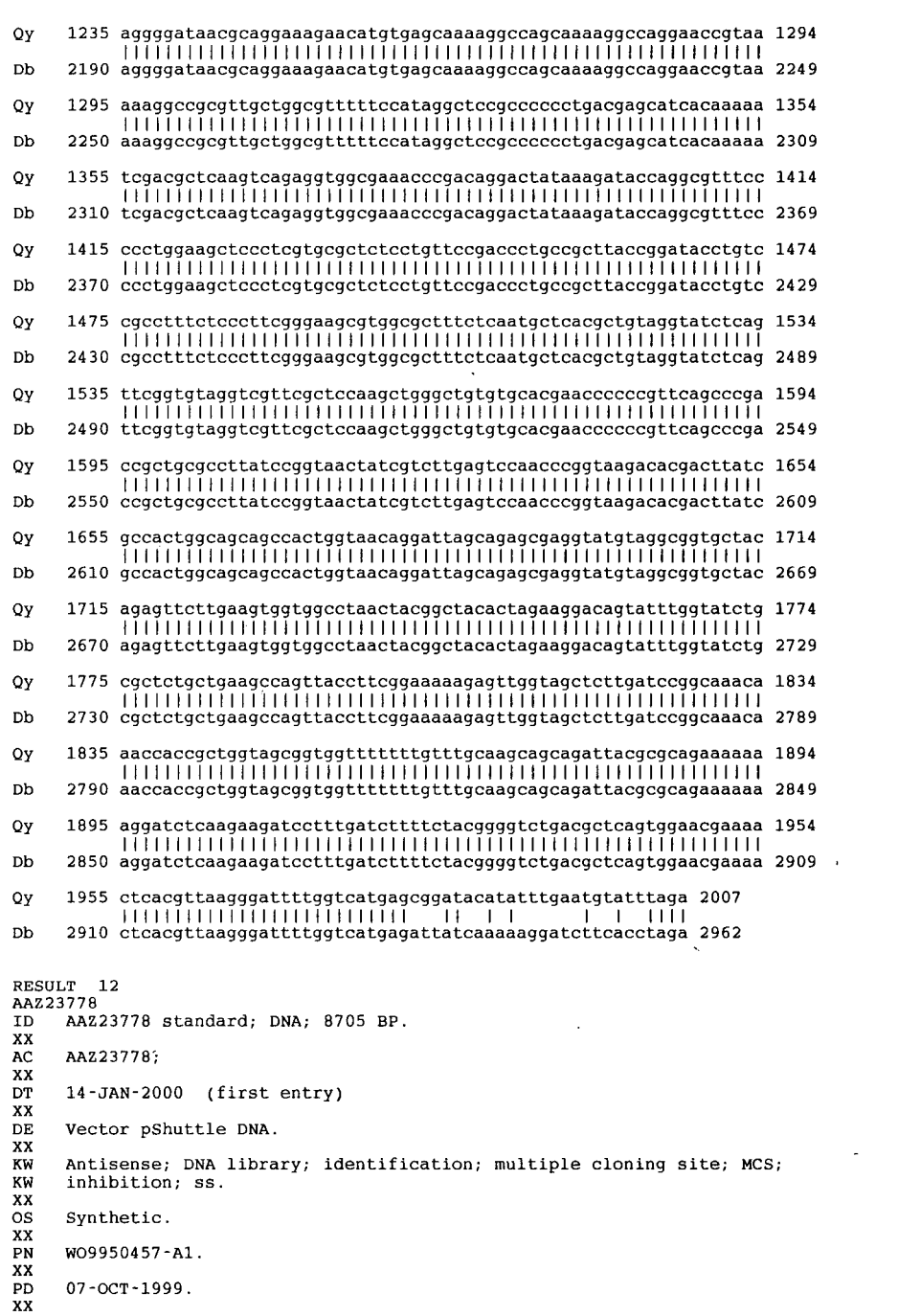
RESULT 9
AAV38297
ID AAV38297 standard; DNA; 5446 BP.
XX
AC AAV38297;
XX
DT 26-OCT-1998 (first entry)
XX
DE Plasmid pcDNA3.
XX
KW Plasmid pcDNA3; pneumococcal surface protein A; PspA; infection;
KW Streptococcus pneumoniae; sepsis; otitis media; meningitis;
KW bacteraemia; pneumonia; vaccine; genetic immunisation; ss.
XX
OS Chimeric - human cytomegalovirus.
XX
PN WO9824927-A1.
XX
PD 11-JUN-1998.
XX
PF 04-DEC-1997; 97WO-US22847.
XX
PR 04-DEC-1996; 96US-0759505.
XX
PA (UYAL-) UNIV ALABAMA.
XX
PI Briles DE, Curiel DT, McDaniel LS;
XX
DR WPI; 1998-333343/29.
XX
PT Plasmid containing pneumococcal epitope for expression in eukaryotic
PT cells - useful for eliciting immunological response to pneumococcal
PT infection or sepsis
XX
PS Example 1; Fig 1B1 to IB-5; 47pp; English.
XX
CC This is the DNA sequence of plasmid pcDNA3 (Invitrogen). A portion
CC of the gene (see AAV38298) that codes for respiratory syncytial virus
CC glycoprotein G (RSVG) has been ampliifed, digested with KpnI and
CC ligated into KpnI-digested pcDNA3 upstream of the multiple cloning
CC site of pcDNA3 and downstream of the human cytomegalovirus
CC immediate early (HCMV-IE) promoter to create pGT41. A full-length
CC coding sequence of Streptococcus pneumoniae Rx1 pneumococcal
CC surface protein A (PspA) was then inserted into pGT41 to create
CC a fusion between rsvG and PspA. Intramuscular immunisation of
CC BALB/c mice with the resulting plasmid, designated pKSD2601,
CC induced protection against an otherwise lethal challenge with a
CC capsular type 3 pneumococcus. A plasmid for expression of

```
CC   pneumococcal epitope DNA in eukaryotic cells is claimed.  The
CC   plasmid includes a promoter for driving expression in a eukaryotic
CC   cell (e.g. HCMV-IE), DNA encoding a leader sequence (e.g. of RSVG)
CC   which facilitates expression, translation through or transport of
CC   the expression product in a eukaryotic cell membrane, and DNA
CC   encoding a pneumonoccal epitope such as PspA.  The invention also
CC   provides a vaccine comprising the plasmid and a suitable carrier or
CC   diluent, and optionally one or more cytokines or DNA encoding them,
CC   or a bacterial delivery system.  The vaccine is used to elicit an
CC   immunological response in a host, including humans, susceptible to
CC   pneumococcal infection or sepsis.  The plasmid can also be used to
CC   express a pneumococcal epitope of interest in vitro.
XX
SQ   Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 other;


  Query Match             13.3%;  Score 1082;  DB 19;  Length 5446;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-153;
  Matches 1082;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     210 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 269

Qy      61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 329

Qy     121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     330 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 389

Qy     181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 449

Qy     241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     450 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 509

Qy     301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     510 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 569

Qy     361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     570 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 629

Qy     421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     630 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 689

Qy     481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     690 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 749

Qy     541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     750 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 809

Qy     601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     810 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 869

Qy     661 actcactatagggagacccaagcttggtaccgagctcggatccactagtaacggccgcca 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     870 actcactatagggagacccaagcttggtaccgagctcggatccactagtaacggccgcca 929

Qy     721 gtgtgctggaattctgcagatatccatcacactggcggccgctcgagcatgcatctagag 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     930 gtgtgctggaattctgcagatatccatcacactggcggccgctcgagcatgcatctagag 989

Qy     781 ggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgactgtgcct 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     990 ggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgactgtgcct 1049

Qy     841 tctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggt 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1050 tctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaaggt 1109

Qy     901 gccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtagg 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1110 gccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtagg 1169

Qy     961 tgtcattctattctgggggggtgggggtgggcaggacagcaagggggaggattgggaagac 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1170 tgtcattctattctgggggggtgggggtgggcaggacagcaagggggaggattgggaagac 1229

Qy    1021 aatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccagc 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1230 aatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccagc 1289

Qy    1081 tg 1082
           ||
Db    1290 tg 1291


RESULT  10
AAA53871
ID   AAA53871 standard; DNA; 5245 BP.
XX
AC   AAA53871;
XX
DT   03-JAN-2001  (first entry)
XX
DE   Expression vector pRIG22b.
XX
KW   Vector; endogenous gene; activation; over-expression;
KW   erythropoietin; growth hormone; drug discovery;
KW   granulocyte colony stimulating factor; ds.
XX
OS   Synthetic.
XX
PN   WO200049162-A2.
XX
PD   24-AUG-2000.
XX
PF   22-FEB-2000; 2000WO-US04429.
XX
PR   19-FEB-1999;   99US-0253022.
PR   08-MAR-1999;   99US-0263814.
PR   26-MAR-1999;   99US-0276820.
XX
PA   (ATHE-) ATHERSYS INC.
XX
PI   Harrington JJ,  Sherf B,  Rundlett S;
XX
DR   WPI; 2000-549276/50.
XX
PT   Non-targeted activation of endogenous genes, e.g. for the production of
PT   erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT   proteins and for drug discovery
XX
PS   Disclosure; Fig 16; 240pp; English.
XX
CC   New methods, vectors and cells are described for non-targeted
CC   activation and over-expression of endogenous genes. The vector
CC   constructs comprise transcriptional regulatory sequences (TRS) and
CC   unpaired splice donor sequences (USDS), preferably the vectors
CC   comprise (in sequential order) a TRS, an USDS, a rare cutting
CC   restriction site (RCRS) and a linearization site (LS) with a second
CC   TRS linked to a selectable marker (SM) lacking a polyadenylation
CC   signal. The methods, vectors and cells comprising the vectors may
CC   be used for the non-targeted activation and over-expression of
CC   endogenous genes, e.g. for the production of proteins (including
CC   erythropoietin, growth hormone or granulocyte-colony stimulating
```

```
CC   factor) and drug discovery. The advantage of these methods are that
CC   endogenous genes including those associated with human disease and
CC   development, may be activated and isolated without prior knowledge
CC   of the sequence structure, function or expression profile of
CC   the genes being known.
XX
SQ   Sequence 5245 BP; 1287 A; 1337 C; 1367 G; 1253 T; 1 other;


  Query Match            13.2%;  Score 1071.6;  DB 21;  Length 5245;
  Best Local Similarity  75.4%;  Pred. No. 8.9e-152;
  Matches 1503;  Conservative    0;  Mismatches  349;  Indels  142;  Gaps    8;

Qy      23 gttgacattgattattgactagttattaatagtaatcaattacggggtcattagttcata 82
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     140 gttggcattgattattgactagttattaatagtaatcaattacggggtcattagttcata 199

Qy      83 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 142
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     200 gcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccgc 259

Qy     143 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     260 ccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatag 319

Qy     203 ggactttccattgacgtcaatgggtggactatttacggtaaactgcccacttggcagtac 262
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     320 ggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtac 379

Qy     263 atcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccg 322
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     380 atcaagtgtatcatatgccaagtccgccccctattgacgtcaatgacggtaaatggcccg 439

Qy     323 cctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacg 382
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     440 cctggcattatgcccagtacatgaccttacgggactttcctacttggcagtacatctacg 499

Qy     383 tattagtcatcgctattaccatggtgatgc-ggttttggcagtacatcaatgggcgtgga 441
           |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||
Db     500 tattagtcatcgctattaccatggtgatgcgggttttggcagtacaccaatgggcgtgga 559

Qy     442 tagcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttg 501
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     560 tagcggtttgactcacggggatttccaagtctccaccccattgacgtcaatgggagtttg 619

Qy     502 ttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaact--------ccgccc 553
           |||||||||||||||||||||||||||||||||||||||||||||        ||||||
Db     620 ttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactgcgatcgcccgccc 679

Qy     554 cattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctct 613
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db     680 cgttgacgcaaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctcgtt 739

Qy     614 ggctaactagagaacccactgcttactggcttatcgaaattaatacgactcactataggg 673
             | |   |  |   ||||     ||   ||   | | || ||   | | |
Db     740 tagtgaaccgtcagatcactagaagctttattgcggtagtttatcacagtta-------- 791

Qy     674 agacccaagcttggtaccgagctcggatccactagtaacggccgccagtgtgctggaatt 733
           |    | | |   ||       || ||  ||  |||  ||  |   ||       || |
Db     792 aattgctaacgcagtcagtgcttctgacacaacagtctcgaacttaagctgcagtgactc 851

Qy     734 ctgcagatatccatcacactggcggccgctcgagcatgcatctagagggccctattctat 793
             | ||   |           |   |  |  ||   | |      ||||    |
Db     852 tcttaaatccaccatggctacaggtgagtactcgctaccttaagagaggcctatctggcc 911

Qy     794 agtgtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagc 853
           |||      |  | |  |||  |    ||    |       | ||     ||     ||
Db     912 agttagcagtcgaagaaagaagtttaagagagccgaaacaagcgctcatgagcccgaagt 971

Qy     854 catctgttgtttgcccctcccccgtgccttccttgaccctggaaggtgccactcccactg 913
             |     |   || |  | |||   ||   ||     ||     || ||  |  |||
Db     972 ggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacctg 1031

Qy     914 tcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattc 973
           |
Db    1032 t---------------------------------------------------------- 1032

Qy     974 tggggggtgggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatg 1033
              ||    ||||  || || ||  | | |   ||  | |     | | || | |  ||
Db    1033 ----ggcgccggtgatgccggccacgatgcgtccggcgtagaggatccacaggacgggtg 1088

Qy    1034 ctggggatgcggtgggctctatggcttctgaggcggaaagaaccagctgcattaatgaat 1093
             |  |    | |                                  ||    |  ||
Db    1089 tggtcgccatgat-------------------------------cgcgtagtcgatagtg 1117

Qy    1094 cggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcac 1153
              ||||   ||| | ||| || || |    || |            |   ||||
Db    1118 gctccaagtagcgaagcgagcaggactgggcggcggccaaagcggtcggacagtgctccg 1177

Qy    1154 tgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggt 1213
            ||   | |||||   |   ||        ||      |    || ||  |   ||
Db    1178 agaacgggtgcgcatagaaattgcatcaacgcatatagcgctagatccttgctagagtcg 1237

Qy    1214 aatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggcca 1273
             |   | ||    |                                 ||||||||||||||||
Db    1238 agatctgtcgagc-------------------------------catgtgagcaaaaggcca 1268

Qy    1274 gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccc 1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1269 gcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccc 1328

Qy    1334 ccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact 1393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1329 ccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggact 1388

Qy    1394 ataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccct 1453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1389 ataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccct 1448

Qy    1454 gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatg 1513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db    1449 gccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatag 1508

Qy    1514 ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgca 1573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1509 ctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgca 1568

Qy    1574 cgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaa 1633
           |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1569 cgaa-ccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaa 1627

Qy    1634 cccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagc 1693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1628 cccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagc 1687

Qy    1694 gaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactag 1753
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1688 ga-gtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactag 1746

Qy    1754 aaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttgg 1813
           ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1747 aaggacactatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttgg 1806

Qy    1814 tagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagca 1873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1807 tagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagca 1866

Qy    1874 gcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtc 1933
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    1867 gcagattacgcggagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtc 1926

Qy    1934 tgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacatatt 1993
           |||||||||||||||||||||||||||||||||||||||||||||||||   ||  | |
Db    1927 tgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaag 1986

Qy    1994 tgaatgtatttaga 2007
```

```
            |  |  ||||
Db   1987 gatcttcacctaga 2000


RESULT 11
AAV40007
ID   AAV40007 standard; DNA; 4026 BP.
XX
AC   AAV40007;
XX
DT   15-FEB-1999  (first entry)
XX
DE   Plasmid pCTMI.
XX
KW   E2F; transcription factor; human; retinoblastoma protein RB;
KW   bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW   thyroid hyperplasia; Grave's disease; psoriasis;
KW   benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW   peripheral vascular disease; therapy; plasmid pCTMI; ss.
XX
OS   Chimeric - cytomegalovirus.
OS   Chimeric - mastadenovirus.
OS   Chimeric - bacteriophage T7.
OS   Chimeric - bacteriophage SP6.
OS   Chimeric - rhesus macaque polyoma virus.
OS   Chimeric - Bos taurus.
XX
FH   Key             Location/Qualifiers
FT   promoter        209..864
FT                   /*tag=  a
FT                   /note= "CMV promoter"
FT   misc_feature    907..1074
FT                   /*tag=  b
FT                   /function= tripartite leader sequence
FT   intron          1075..1253
FT                   /*tag=  c
FT                   /note= "hybrid SV40 late intron"
FT   promoter        1305..1322
FT                   /*tag=  d
FT                   /note= "SP6 promoter"
FT   misc_feature    1851..4026
FT                   /*tag=  e
FT                   /note= "pUC19 backbone H3 to AatII"
FT   CDS             complement (3032..3890)
FT                   /*tag=  f
FT                   /note= "AMP-ORF"
XX
PN   WO9821228-A1.
XX
PD   22-MAY-1998.
XX
PF   13-NOV-1997;   97WO-US21821.
XX
PR   14-FEB-1997;   97US-0801092.
PR   15-NOV-1996;   96US-0751517.
XX
PA   (CANJ-) CANJI INC.
XX
PI   Antelman D,  Gregory RJ,  Wills KN;
XX
DR   WPI; 1998-297858/26.
XX
PT   New fusion polypeptide of, e.g. transcription factor  - used to
PT   treat, e.g. hyper-proliferative disease such as cancer and
PT   restenosis
XX
PS   Example 1; Fig 6; 91pp; English.
XX
CC   This is the nucleotide sequence of pCTMI, a plasmid that was
CC   constructed from pCTM (see AAV40006) by digesting pCTM with XhoI and
CC   NotI and subcloning a 180 bp intron XhoI-NotI fragment from a
CC   pCMV-beta-gal vector.  Plasmid pCTMI has been used as a vector for
CC   the expression of fusion proteins of the invention that comprise
CC   retinoblastoma protein (BP, see AAW62465) and E2F transcription
CC   factor (see AAW62464).  Such fusion proteins, particularly expressed
CC   from gene therapy vectors, are used to treat hyperproliferative
CC   conditions, specifically cancer (particularly of the bladder) or
CC   restenosis.  They are more effective in repressing transcription of
CC   the E2F promoter than RB alone and cause cell-cycle arrest in a
CC   variety of cells.
XX
SQ   Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 other;


  Query Match             13.2%;  Score 1070.2;  DB 19;  Length 4026;
  Best Local Similarity   69.7%;  Pred. No. 1.5e-151;
  Matches 1919;  Conservative    0;  Mismatches   88;  Indels  746;  Gaps    5;

Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    210 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 269

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    270 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 329

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    330 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 389

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    390 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 449

Qy    241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    450 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 509

Qy    301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    510 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 569

Qy    361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    570 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 629

Qy    421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    630 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 689

Qy    481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    690 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 749

Qy    541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    750 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 809

Qy    601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    810 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 869

Qy    661 actcactatagggagacccaagcttgg---------------------------------- 687
          |||||||||||||||||||||||||| |
Db    870 actcactatagggagacccaagcttcgcgcgggtaccactctcttccgcatcgctgtctg 929

Qy    688 ---------------------------------------------------------- 687

Db    930 cgagggccagctgttgggctcgcggttgaggacaaactcttcgcggtctttccagtactc 989

Qy    688 ---------------------------------------------------------- 687

Db    990 ttggatcggaaacccgtcggcctccgaacggtactccgccaccgagggacctgagcgagt 1049

Qy    688 ---------------------------------------------------------- 687
```





RESULT 12
AAZ23778
ID AAZ23778 standard; DNA; 8705 BP.
XX
AC AAZ23778;
XX
DT 14-JAN-2000 (first entry)
XX
DE Vector pShuttle DNA.
XX
KW Antisense; DNA library; identification; multiple cloning site; MCS;
KW inhibition; ss.
XX
OS Synthetic.
XX
PN WO9950457-A1.
XX
PD 07-OCT-1999.
XX

```
PF   28-MAR-1999;   99WO-US06742.
XX
PR   28-MAR-1998;   98US-0079792.
PR   06-NOV-1998;   98US-0107504.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
XX
PI   Ruffner DE,  Pierce ML,  Chen Z;
XX
DR   WPI; 1999-610866/52.
XX
PT   Production of antisense libraries, used for identifying antisense
PT   agents and for identifying target sites for antisense-mediated
PT   inhibition of a selected gene  -
XX
PS   Claim 16; Page 43-50; 63pp; English.
XX
CC   This invention describes a novel method for generating an antisense
CC   library targeted to a selected RNA transcript. The methods can be used
CC   for identifying antisense agents and for identifying target sites for
CC   antisense-mediated inhibition of a selected gene. The use of a direct
CC   library for target site selection significantly simplifies the screening
CC   process, since only very small libraries need be prepared and assayed.
CC   This sequence represents the vector pShuttle which is used in the method
CC   of the invention.
XX
SQ   Sequence 8705 BP; 1875 A; 2575 C; 2078 G; 2177 T; 0 other;


  Query Match           13.2%;  Score 1068;  DB 20;  Length 8705;
  Best Local Similarity  99.9%;  Pred. No. 2.9e-151;
  Matches 1079;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2708 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 2767

Qy     61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2768 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 2827

Qy    121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2828 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 2887

Qy    181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2888 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 2947

Qy    241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2948 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 3007

Qy    301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3008 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 3067

Qy    361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3068 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 3127

Qy    421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3128 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 3187

Qy    481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3188 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 3247

Qy    541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3248 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 3307

Qy     601 agcagagctctctggctaactagagaacccactgctt-actggcttatcgaaattaatac 659
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    3308 agcagagctctctggctaactagagaacccactgcttaactggcttatcgaaattaatac 3367

Qy     660 gactcactatagggagacccaagcttggtaccgagctcggatccactagtaacggccgcc 719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3368 gactcactatagggagacccaagcttggtaccgagctcggatccactagtaacggccgcc 3427

Qy     720 agtgtgctggaattctgcagatatccatcacactggcggccgctcgagcatgcatctaga 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3428 agtgtgctggaattctgcagatatccatcacactggcggccgctcgagcatgcatctaga 3487

Qy     780 gggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgactgtgcc 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3488 gggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgactgtgcc 3547

Qy     840 ttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaagg 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3548 ttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaagg 3607

Qy     900 tgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtag 959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3608 tgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtag 3667

Qy     960 gtgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaaga 1019
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3668 gtgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaaga 3727

Qy    1020 caatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccag 1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3728 caatagcaggcatgctggggatgcggtgggctctatggcttctgaggcggaaagaaccag 3787


RESULT  13
AAA47797
ID   AAA47797 standard; DNA; 6050 BP.
XX
AC   AAA47797;
XX
DT   08-NOV-2000  (first entry)
XX
DE   pWRG/SEO-S vaccine construct.
XX
KW   Seoul virus; hantavirus; vaccine; humoral immune response;
KW   cellular immune response; detection; pathogen; immunity; ds.
XX
OS   Synthetic.
XX
PN   WO200044406-A2.
XX
PD   03-AUG-2000.
XX
PF   27-JAN-2000; 2000WO-US01999.
XX
PR   29-JAN-1999;   99US-0117680.
XX
PA   (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI   Schmaljohn CS,  Hooper JW;
XX
DR   WPI; 2000-491219/43.
XX
PT   Carrier particle is coated with a DNA sequence comprising a promoter
PT   and protein coding region for a determinant of a hantavirus protein,
PT   used as a vaccine against a hantavirus infection
XX
PS   Example 1; Page 58-61; 64pp; English.
XX
CC   New compositions are described in which a carrier particle is coated
CC   with a DNA sequence comprising a promoter operative in the cells of a
CC   mammal and a protein coding region coding for a determinant of a
CC   hantavirus protein. The composition is used as a vaccine against
CC   hantavirus infection. Intracellular expression of the hantavirus
```

```
CC   genome segments elicits humoral and cell-mediated immune responses
CC   against hantavirus. The DNA sequences may also be used for detection
CC   of the virus. The multivalent vaccine protects against pathologies
CC   resulting from exposure to one or several hantaviruses. The
CC   advantages of the composition are that the risks of pathogenic
CC   infection are not present as live vaccines are not used and
CC   production of the composition does not require growth or use of
CC   hantavirus which can be spread by aerosol transmission. To
CC   determine the efficacy of the S gene segment of Seoul virus to
CC   provide protective immunity the gene segment was subcloned into the
CC   naked DNA expression vector pWRG7077 downstream of the
CC   cytomegalovirus immediate early promoter to yield pWRG/SEO-S. The
CC   construct was then used to challenge hamsters. Postchallenge serum
CC   samples from pWRG/SEO-S vaccinated hamsters were positive for
CC   glycoprotein and nucleocapsid antibodies.
XX
SQ   Sequence 6050 BP; 1634 A; 1416 C; 1454 G; 1546 T; 0 other;


  Query Match             13.0%;  Score 1058.4;  DB 21;  Length 6050;
  Best Local Similarity   94.4%;  Pred. No. 8.1e-150;
  Matches 1121;  Conservative    0;  Mismatches   36;  Indels   31;  Gaps    1;

Qy     820 tgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtg 879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4754 tgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtg 4813

Qy     880 ccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaatt 939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4814 ccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaatt 4873

Qy     940 gcatcgcattgtctgagtaggtgtcattctattctgggggtggggtggggcaggacagc 999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4874 gcatcgcattgtctgagtaggtgtcattctattctgggggtggggtggggcaggacagc 4933

Qy    1000 aaggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggct 1059
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4934 aaggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggct 4993

Qy    1060 tctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119
           ||||||||||||||||||||||                              || | |
Db    4994 tctgaggcggaaagaaccagct------------------------------ggggctc 5022

Qy    1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179
             |   | |           ||||||||||||||||||||||||||||||||||||||||||
Db    5023 gacagctcgactctagaattgcttcctcgctcactgactcgctgcgctcggtcgttcggc 5082

Qy    1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5083 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 5142

Qy    1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5143 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 5202

Qy    1300 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 1359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5203 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 5262

Qy    1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 1419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5263 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 5322

Qy    1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5323 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 5382

Qy    1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5383 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 5442

Qy    1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgct 1599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5443 tgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgct 5502

Qy    1600 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5503 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 5562

Qy    1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5563 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 5622

Qy    1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5623 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 5682

Qy    1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5683 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 5742

Qy    1840 ccgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5743 ccgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 5802

Qy    1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5803 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 5862

Qy    1960 gttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
           |||||||||||||||||||   ||  | |      |  |  ||||
Db    5863 gttaagggattttggtcatgagattatcaaaaaggatcttcacctaga 5910


RESULT  14
AAA47796
ID   AAA47796 standard; DNA; 8002 BP.
XX
AC   AAA47796;
XX
DT   08-NOV-2000  (first entry)
XX
DE   pWRG/SEO-M vaccine construct.
XX
KW   Seoul virus; hantavirus; vaccine; humoral immune response;
KW   cellular immune response; detection; pathogen; immunity; ds.
XX
OS   Synthetic.
XX
PN   WO200044406-A2.
XX
PD   03-AUG-2000.
XX
PF   27-JAN-2000; 2000WO-US01999.
XX
PR   29-JAN-1999;   99US-0117680.
XX
PA   (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI   Schmaljohn CS,  Hooper JW;
XX
DR   WPI; 2000-491219/43.
XX
PT   Carrier particle is coated with a DNA sequence comprising a promoter
PT   and protein coding region for a determinant of a hantavirus protein,
PT   used as a vaccine against a hantavirus infection
XX
PS   Example 1; Page 55-58; 64pp; English.
XX
CC   New compositions are described in which a carrier particle is coated
CC   with a DNA sequence comprising a promoter operative in the cells of a
CC   mammal and a protein coding region coding for a determinant of a
CC   hantavirus protein. The composition is used as a vaccine against
CC   hantavirus infection. Intracellular expression of the hantavirus
CC   genome segments elicits humoral and cell-mediated immune responses
CC   against hantavirus. The DNA sequences may also be used for detection
```

```
CC   of the virus. The multivalent vaccine protects against pathologies
CC   resulting from exposure to one or several hantaviruses. The
CC   advantages of the composition are that the risks of pathogenic
CC   infection are not present as live vaccines are not used and
CC   production of the composition does not require growth or use of
CC   hantavirus which can be spread by aerosol transmission. To
CC   determine the efficacy of the M gene segment of Seoul virus to
CC   provide protective immunity the gene segment was subcloned into the
CC   naked DNA expression vector pWRG7077 downstream of the
CC   cytomegalovirus immediate early promoter to yield pWRG/SEO-M. The
CC   construct was then used to challenge 6 to 8 week old female BALB/c
CC   mice. Mice vaccinated with pWRG/SEO-M produced neutralising
CC   antibodies.
XX
SQ   Sequence 8002 BP; 2208 A; 1788 C; 1818 G; 2188 T; 0 other;


  Query Match            13.0%;  Score 1058.4;  DB 21;  Length 8002;
  Best Local Similarity  94.4%;  Pred. No. 7.8e-150;
  Matches 1121;  Conservative    0;  Mismatches   36;  Indels   31;  Gaps    1;

Qy     820 tgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtg 879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6706 tgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtg 6765

Qy     880 ccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaatt 939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6766 ccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaatt 6825

Qy     940 gcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagc 999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6826 gcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagc 6885

Qy    1000 aaggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggct 1059
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6886 aaggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggct 6945

Qy    1060 tctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119
           ||||||||||||||||||||                                  || | |
Db    6946 tctgaggcggaaagaaccagct------------------------------ggggctc 6974

Qy    1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179
             |   | |           ||||||||||||||||||||||||||||||||||||||||||
Db    6975 gacagctcgactctagaattgcttcctcgctcactgactcgctgcgctcggtcgttcggc 7034

Qy    1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7035 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 7094

Qy    1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7095 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 7154

Qy    1300 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 1359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7155 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 7214

Qy    1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctg 1419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7215 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctg 7274

Qy    1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7275 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 7334

Qy    1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7335 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 7394

Qy    1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgct 1599
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7395 tgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgct 7454

Qy    1600 gcgcctttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7455 gcgcctttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 7514

Qy    1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7515 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 7574

Qy    1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7575 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 7634

Qy    1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7635 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 7694

Qy    1840 ccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7695 ccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 7754

Qy    1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7755 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 7814

Qy    1960 gttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
           |||||||||||||||||||    ||  | |      |  |  ||||
Db    7815 gttaagggattttggtcatgagattatcaaaaaggatcttcacctaga 7862


RESULT 15
AAA27831
ID   AAA27831 standard; DNA; 3796 BP.
XX
AC   AAA27831;
XX
DT   12-SEP-2000  (first entry)
XX
DE   Vector plasmid pCMV-MC1.
XX
KW   North American PRRS virus; Nidovirales virus; pig; swine; vaccine;
KW   pCMV-S-P129; ss.
XX
OS   Chimeric - Porcine reproductive and respiratory syndrome virus.
OS   Chimeric - Human cytomegalovirus.
XX
PN   EP1018557-A2.
XX
PD   12-JUL-2000.
XX
PF   25-NOV-1999;   99EP-0309409.
XX
PR   22-DEC-1998;   98US-0113345.
XX
PA   (PFIZ ) PFIZER PROD INC.
XX
PI   Calvert JG,  Welch SW,  Sheppard MG;
XX
DR   WPI; 2000-444364/39.
XX
PT   New polynucleotide encoding an infectious RNA molecule of a North
PT   American porcine reproductive and respiratory syndrome virus for use as
PT   a vaccine in protecting swine and other animals from infection by a
PT   pathogen -
XX
PS   Example 4; Page 43-44; 53pp; English.
XX
CC   The present sequence is that of eukaryotic expression vector
CC   plasmid pCMV-MC1.  The plasmid is derived from pCMVbeta by
CC   replacing the LacZ coding sequence with a linker containing
CC   multiple restriction sites.  Modification of the human
CC   cytomegalovirus (CMV) immediate early promoter was accomplished by
CC   substituting the sequence between SacI and the 2nd NotI site of
CC   pCMV-MC1 with a synthetic linker.  North American porcine
```

```
CC   reproductive and respiratory syndrome (PRRS) virus P129A cDNA (see
CC   AAA27809) was cloned into the modified vector to create pCMV-S-P129
CC   (ATCC 203489), which was used to demonstrate cellular expression of
CC   PRRS virus by direct transfection of cDNA into cells.  The
CC   invention relates to polynucleotide molecules, plasmids, viral
CC   vectors and transfected host cells that comprise North American
CC   PRRS DNA.  It also relates to polynucleotide molecules, viral
CC   vectors and transfected host cells encoding a genetically modified
CC   North American PRRS virus that is disabled in its ability to cause
CC   PRRS, or which encodes 1 or more heterologous antigenic epitopes,
CC   for use as a vaccine.
XX
SQ   Sequence 3796 BP; 953 A; 934 C; 956 G; 953 T; 0 other;

  Query Match             12.8%;  Score 1036.2;  DB 21;  Length 3796;
  Best Local Similarity   71.8%;  Pred. No. 1.7e-146;
  Matches 1613;  Conservative    0;  Mismatches  293;  Indels  341;  Gaps    6;

Qy    102 cgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccatt 161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 cgttacataacttacggtaaatggcccgcctggctgaccgcccaacgacccccgcccatt  86

Qy    162 gacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtca 221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 gacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtca 146

Qy    222 atgggtggactatttacggtaaactgcccacttggcagtacatcaagtgtatcatatgcc 281
          |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 atgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatcatatgcc 206

Qy    282 aagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagta 341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 aagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagta 266

Qy    342 catgaccttatgggactttcctacttggcagtacatctacgtattagtcatcgctattac 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 catgaccttatgggactttcctacttggcagtacatctacgtattagtcatcgctattac 326

Qy    402 catggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacgggg 461
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 catggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacgggg 386

Qy    462 atttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacg 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 atttccaagtctccaccccattgacgtcaatgggagtttgttttggcaccaaaatcaacg 446

Qy    522 ggactttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgt 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 ggactttccaaaatgtcgtaacaactccgccccattgacgcaaatgggcggtaggcgtgt 506

Qy    582 acggtgggaggtctatataagcagagctctctggctaa---------------------- 619
          ||||||||||||||||||||||||||||||  |   | |
Db    507 acggtgggaggtctatataagcagagctcgtttagtgaaccgtcagatcgcctggagacg 566

Qy    620 ----------------------------------------------------------ct 621
                                                                     ||
Db    567 ccatccacgctgttttgacctccatagaagacaccgggaccgatccagcctccggactct 626

Qy    622 agagaacccactgct-----tactggcttatcgaaattaatacgactcactatagggaga 676
          |||| | ||  | ||      ||||    | |  ||     ||   | | | |||
Db    627 agaggatccggtactcgaggaactgaaaaaccagaaagttaactggtaagtttagtcttt 686

Qy    677 cccaagcttggtaccgagctcggatccactagtaacggccgccagtgtgctggaattctg 736
                ||  | | |  | |||||||  | ||   |     ||  |  |||    || |
Db    687 ttgtcttttatttcaggtcccggatccggtggtggtgcaaatcaaagaactgctcctcag 746

Qy    737 cagatatccatcacactggcggccgctcgagcatgcatctagagggccctattctatagt 796
            ||| |       |||    | |      | | |  |      || |||     |
Db    747 tggatgttgcctttacttctaggcctgtacggaagtgttacttctgctctaaaagctgcg 806

Qy    797 gtcacctaaatgctagagctcgctgatcagcctcgactgtgccttctagttgccagccat 856
          |     ||   ||    ||  | |      |   ||   |  |      |||  | | ||

Db    807 gaattgtacccgcggccgcaagatatcgccctaggaagatctcgatcgattggtaccaat 866

Qy    857 ctg--------ttgtttgccccctcccccgtgccttccttgaccctggaaggtgccactcc 908
          | |        |  ||  |  ||  |   |       |  | ||| |||   |      |
Db    867 ccgcgacccctttaattaacagctagcggatttaaatcagggcccgggatactagtgagcg 926

Qy    909 cactgtccttttcctaataaaatgaggaaattgcatcgcattgtctga------------- 955
            |       |||  | |  || ||  |  |  ||    |||    |
Db    927 gccgcggggatccagacatgataagatacattgatgagtttggacaaaccacaactagaa 986

Qy    956 ---------------------------------------------------------- 955

Db    987 tgcagtgaaaaaaatgctttatttgtgaaatttgtgatgctattgctttatttgtaacca 1046

Qy    956 ------------------------------------gtaggtgtcattctattctggggg 979
                                               | |    |  |    ||| ||
Db   1047 ttataagctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttc 1106

Qy    980 gtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggg- 1038
            |||| || |  |||         |||   |   ||    |   ||||||||||  |
Db   1107 agggggaggtgtgggaggttttttcggatcctctagagtcgacctgcaggcatgcaagct 1166

Qy   1039 ---------------------------------------------------------- 1038

Db   1167 tggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccac 1226

Qy   1039 ---------------------------------------------------------- 1038

Db   1227 acaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaac 1286

Qy   1039 --------gatgcggtgggctctatggcttctgaggcggaaagaa----------ccagc 1080
                   |||| || ||||  || |  ||    |   |||          |||||
Db   1287 tcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagc 1346

Qy   1081 tgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1347 tgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccg 1406

Qy   1141 cttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctc 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 cttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctc 1466

Qy   1201 actcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgt 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 actcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgt 1526

Qy   1261 gagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttcc 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 gagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttcc 1586

Qy   1321 ataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaa 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1587 ataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaa 1646

Qy   1381 acccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctc 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 acccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctc 1706

Qy   1441 ctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtgg 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 ctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtgg 1766

Qy   1501 cgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagc 1560
          ||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1767 cgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagc 1826

Qy   1561 tgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatc 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1827 tgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatc 1886

Qy   1621 gtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaaca 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1887 gtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaaca 1946
```

```
Qy   1681 ggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaact 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1947 ggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaact 2006

Qy   1741 acggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcg 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2007 acggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcg 2066

Qy   1801 gaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttt 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2067 gaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttt 2126

Qy   1861 ttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatct 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2127 ttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatct 2186

Qy   1921 tttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2187 tttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatga 2246

Qy   1981 gcggatacatatttgaatgtatttaga 2007
          |   ||  | |      |  |  ||||
Db   2247 gattatcaaaaaggatcttcacctaga 2273
```

Search completed: November 3, 2001, 20:43:59
Job time: 16975 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:49:24 ; Search time 154.79 Seconds
(without alignments)
9924.800 Million cell updates/sec

Title: US-09-778-516-1
Perfect score: 8115
Sequence: 1 gatgtacgggccagatatac.....gcgtgtcgctagcgacggtc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1240.4	15.3	3853	3	US-08-801-092-5	Sequence 5, Appli
2	1068.6	13.2	4026	3	US-08-801-092-19	Sequence 19, Appl
3	1053.6	13.0	4326	4	US-08-760-615-7	Sequence 7, Appli
4	981.8	12.1	6253	3	US-08-893-327-15	Sequence 15, Appl
5	981.8	12.1	6280	3	US-08-893-327-17	Sequence 17, Appl
6	981.8	12.1	6280	3	US-08-893-327-19	Sequence 19, Appl
7	973	12.0	4283	1	US-08-343-401A-3	Sequence 3, Appli
8	973	12.0	4283	1	US-08-445-265A-1	Sequence 1, Appli
9	973	12.0	4283	3	US-08-990-442-1	Sequence 1, Appli
10	921.8	11.4	6285	1	US-08-467-420A-49	Sequence 49, Appl
11	921.8	11.4	6285	1	US-08-470-110A-49	Sequence 49, Appl
12	921.8	11.4	6285	1	US-08-667-769A-49	Sequence 49, Appl
13	921.8	11.4	6285	2	US-08-940-371-49	Sequence 49, Appl
14	921.8	11.4	6285	3	US-08-637-647-49	Sequence 49, Appl
15	921.8	11.4	6285	5	PCT-US95-17082A-49	Sequence 49, Appl
16	906.4	11.2	3130	4	US-09-038-141-1	Sequence 1, Appli
17	906.4	11.2	3789	4	US-09-075-019-8	Sequence 8, Appli
18	906.4	11.2	3918	2	US-08-495-500-2	Sequence 2, Appli
19	906.4	11.2	4249	3	US-08-801-092-33	Sequence 33, Appl
c 20	906.4	11.2	4453	1	US-08-770-761A-4	Sequence 4, Appli
21	906.4	11.2	4454	2	US-08-738-172-2	Sequence 2, Appli
c 22	906.4	11.2	4525	2	US-08-613-861-2	Sequence 2, Appli
c 23	906.4	11.2	4540	1	US-08-770-761A-6	Sequence 6, Appli
24	906.4	11.2	4824	2	US-08-485-139-5	Sequence 5, Appli
25	906.4	11.2	4824	3	US-08-750-357-5	Sequence 5, Appli
26	906.4	11.2	4883	1	US-08-064-121-4	Sequence 4, Appli
27	906.4	11.2	4883	1	US-08-318-772A-1	Sequence 1, Appli
28	906.4	11.2	4883	1	US-08-478-015-4	Sequence 4, Appli
29	906.4	11.2	4883	3	US-08-475-975-4	Sequence 4, Appli
30	906.4	11.2	4883	3	US-09-084-889-4	Sequence 4, Appli
31	906.4	11.2	4949	3	US-09-138-024-22	Sequence 22, Appl
32	906.4	11.2	5399	1	US-08-064-121-1	Sequence 1, Appli
33	906.4	11.2	5399	1	US-08-478-015-1	Sequence 1, Appli
34	906.4	11.2	5399	3	US-08-475-975-1	Sequence 1, Appli
35	906.4	11.2	5399	3	US-09-084-889-1	Sequence 1, Appli
36	906.4	11.2	5618	3	US-08-799-569-1	Sequence 1, Appli
37	906.4	11.2	5620	1	US-08-104-072B-9	Sequence 9, Appli
38	906.4	11.2	5620	1	US-08-351-413-3	Sequence 3, Appli
39	906.4	11.2	5620	2	US-09-025-583-3	Sequence 3, Appli
40	906.4	11.2	5642	1	US-08-318-772A-2	Sequence 2, Appli
41	906.4	11.2	5653	1	US-08-073-836-3	Sequence 3, Appli
42	906.4	11.2	5653	1	US-08-235-277-1	Sequence 1, Appli
43	906.4	11.2	5703	1	US-08-467-420A-50	Sequence 50, Appl
44	906.4	11.2	5703	1	US-08-470-110A-50	Sequence 50, Appl
45	906.4	11.2	5703	1	US-08-667-769A-50	Sequence 50, Appl

ALIGNMENTS

```
RESULT   1
US-08-801-092-5
; Sequence 5, Application US/08801092
; Patent No. 6074850
;  GENERAL INFORMATION:
;    APPLICANT:  Antelman, Douglas
;    APPLICANT:  Gregory, Richard J.
;    APPLICANT:  Wils, Kenneth N.
;    TITLE OF INVENTION:  Tissue Specific Expression of
;    TITLE OF INVENTION:  Retinoblastoma Protein
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/801,092
;      FILING DATE:  14-FEB-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/751,517
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  016930-001020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  703-576-0300
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3853 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  209..250
;    FEATURE:
;      NAME/KEY:  CDS
```

```
;      LOCATION:  254..289
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  293..505
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  509..514
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  518..520
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  524..658
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  662..691
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  695..748
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  752..781
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  785..829
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1132..1134
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1138..1149
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  833..862
US-08-801-092-5

  Query Match             15.3%;  Score 1240.4;  DB 3;  Length 3853;
  Best Local Similarity   74.1%;  Pred. No. 1e-268;
  Matches 1911;  Conservative    0;  Mismatches   96;  Indels  573;  Gaps    4;

Qy       1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     210 GATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCA 269

Qy      61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270 ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA 329

Qy     121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     330 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT 389

Qy     181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGG 449

Qy     241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     450 TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC 509

Qy     301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     510 GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT 569

Qy     361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     570 CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG 629

Qy     421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     630 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC 689

Qy     481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     690 ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT 749

Qy     541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
           ||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db     750 AACAACTCCGCCCCATTGACGCAAATGGGCGGTAGCGCTGTACGGTGGGAGGTCTATATA 809

Qy     601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     810 AGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACG 869

Qy     661 actcactatagggagacccaagctt---------------------------------- 685
           |||||||||||||||||||||||||
Db     870 ACTCACTATAGGGAGACCCAAGCTTCGCGCGGGTACCACTCTCTTCCGCATCGCTGTCTG 929

Qy     686 ---------------------------------------------------------- 685

Db     930 CGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTC 989

Qy     686 ------------------ggtaccgagctcggatccactagtaacggccgccagtgtgc 726
                             |   |||| |     ||| | |    | || |   || | |
Db     990 TTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGT 1049

Qy     727 tggaattctgcagatatccatcacactggcggccgc------------------tcgag 767
             | ||    | |||    |   | || | ||| ||                  ||| |
Db    1050 CCGCATCGACCGGATCGGAAAACCTCTCGAGGCGGCCGCTGCAGTCTAGACGAATTCGCG 1109

Qy     768 catgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctgatcagc 827
             | |        ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1110 TACGATATCGATGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGC 1169

Qy     828 ctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttcctt 887
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1170 CTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTT 1229

Qy     888 gaccctggaaggtgccactcccactgtccttttcctaataaaatgaggaaattgcatcgca 947
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1230 GACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCA 1289

Qy     948 ttgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaaggggga 1007
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1290 TTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGA 1349

Qy    1008 ggattgggaagacaatagcaggcatgctgggga--------------------------- 1040
           |||||||||||||||||| |  |||  |
Db    1350 GGATTGGGAAGACAATAGCCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAG 1409

Qy    1041 ---------------------------------------------------------- 1040

Db    1410 ATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACG 1469

Qy    1041 ---------------------------------------------------------- 1040

Db    1470 CCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACT 1529

Qy    1041 ---------------------------------------------------------- 1040

Db    1530 TGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATA 1589

Qy    1041 ---------------------------------------------------------- 1040

Db    1590 AAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATC 1649

Qy    1041 ---------------------------------------------------------- 1040

Db    1650 ATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTC 1709

Qy    1041 ---------------------------------------------------------- 1040

Db    1710 CTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGT 1769

Qy    1041 ------------------------------------------tgcggtgggctctatgg 1057
```

```
                                                     ||||  ||  ||||   ||
Db    1770 GTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC 1829

Qy    1058 cttct----------gaggcggaaagaaccagctgcattaatgaatcggccaacgcgcgg 1107
           |  ||          ||  |     |  ||||||||||||||||||||||||||||||||
Db    1830 CCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGG 1889

Qy    1108 ggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgct 1167
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1890 GGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT 1949

Qy    1168 cggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatcca 1227
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1950 CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA 2009

Qy    1228 cagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccagga 1287
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2010 CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGA 2069

Qy    1288 accgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatc 1347
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2070 ACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC 2129

Qy    1348 acaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccagg 1407
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2130 ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGG 2189

Qy    1408 cgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggat 1467
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2190 CGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGAT 2249

Qy    1468 acctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggt 1527
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2250 ACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGT 2309

Qy    1528 atctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttc 1587
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2310 ATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTC 2369

Qy    1588 agcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacg 1647
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2370 AGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG 2429

Qy    1648 acttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcg 1707
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2430 ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG 2489

Qy    1708 gtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttg 1767
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2490 GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTG 2549

Qy    1768 gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccg 1827
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2550 GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG 2609

Qy    1828 gcaaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattacgcgca 1887
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2610 GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA 2669

Qy    1888 gaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtgga 1947
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2670 GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA 2729

Qy    1948 acgaaaactcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
           ||||||||||||||||||||||||||||||||    ||  | |      |  |  ||||
Db    2730 ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 2789


RESULT   2
US-08-801-092-19
; Sequence 19, Application US/08801092
; Patent No. 6074850
;  GENERAL INFORMATION:
;    APPLICANT:  Antelman, Douglas
;    APPLICANT:  Gregory, Richard J.
;    APPLICANT:  Wils, Kenneth N.
;    TITLE OF INVENTION:  Tissue Specific Expression of
;    TITLE OF INVENTION:  Retinoblastoma Protein
;    NUMBER OF SEQUENCES:  46
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND and CREW LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/801,092
;      FILING DATE:  14-FEB-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/751,517
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitts, Renee A.
;      REGISTRATION NUMBER:  35,136
;      REFERENCE/DOCKET NUMBER:  016930-001020
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-576-0200
;      TELEFAX:  703-576-0300
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4026 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  209..250
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  254..289
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  293..505
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  509..514
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  518..520
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  524..658
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  662..691
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  695..748
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  752..781
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  785..829
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  833..862
```

```
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1305..1307
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1311..1322
US-08-801-092-19


  Query Match             13.2%;  Score 1068.6;  DB 3;  Length 4026;
  Best Local Similarity   69.7%;  Pred. No. 3.2e-230;
  Matches 1918;  Conservative    0;  Mismatches   89;  Indels  746;  Gaps    5;

Qy       1 gatgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     210 GATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCA 269

Qy      61 attacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggta 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     270 ATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTA 329

Qy     121 aatggcccgcctggctgaccgcccaacgacccccgcccattgacgtcaataatgacgtat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     330 AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT 389

Qy     181 gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacgg 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGACTATTTACGG 449

Qy     241 taaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccctattgac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     450 TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGAC 509

Qy     301 gtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggacttt 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     510 GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTT 569

Qy     361 cctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttgg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     570 CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG 629

Qy     421 cagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     630 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC 689

Qy     481 attgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgt 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     690 ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT 749

Qy     541 aacaactccgccccattgacgcaaatgggcggtaggcgtgtacggtgggaggtctatata 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     750 AACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA 809

Qy     601 agcagagctctctggctaactagagaacccactgcttactggcttatcgaaattaatacg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     810 AGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACG 869

Qy     661 actcactatagggagacccaagcttgg---------------------------------- 687
           |||||||||||||||||||||||||| |
Db     870 ACTCACTATAGGGAGACCCAAGCTTCGCGCGGGTACCACTCTCTTCCGCATCGCTGTCTG 929

Qy     688 ---------------------------------------------------------- 687

Db     930 CGAGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTC 989

Qy     688 ---------------------------------------------------------- 687

Db     990 TTGGATCGGAAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAGCGAGT 1049

Qy     688 ---------------------------------------------------------- 687

Db    1050 CCGCATCGACCGGATCGGAAAACCTCTCGAGGAACTGAAAAACCAGAAAGTTAACTGGTA 1109

Qy     688 ----------------------taccgagctcggatccactagtaacggccgccagtg 723
                                  | | |  | |||||||  | ||   |      ||  |
Db    1110 AGTTTAGTCTTTTTGTCTTTTTATTTCAGGTCCCGGATCCGGTGGTGGTGCAAATCAAAG 1169

Qy     724 tgctggaattctgcagatat---------------------------------------- 743
             |||    || |  ||| |
Db    1170 AACTGCTCCTCAGTGGATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACTTCTGC 1229

Qy     744 -------------ccatcacactggcggccgc----------------tcgagcatgcat 774
                         ||   ||  ||||||||                ||| | | |
Db    1230 TCTAAAAGCTGCGGAATTGTACCCGCGGCCGCTGCAGTCTAGACGAATTCGCGTACGATA 1289

Qy     775 ctagagggccctattctatagtgtcacctaaatgctagagctcgctgatcagcctcgact 834
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1290 TCGATGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACT 1349

Qy     835 gtgccttctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctg 894
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1350 GTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTG 1409

Qy     895 gaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctg 954
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1410 GAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG 1469

Qy     955 agtaggtgtcattctattctgggggggtggggtggggcaggacagcaaggggggaggattgg 1014
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1470 AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGG 1529

Qy    1015 gaagacaatagcaggcatgctgggga---------------------------------- 1040
           |||||||||||| |  |||   |
Db    1530 GAAGACAATAGCCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGA 1589

Qy    1041 ---------------------------------------------------------- 1040

Db    1590 TTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTG 1649

Qy    1041 ---------------------------------------------------------- 1040

Db    1650 GATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTAT 1709

Qy    1041 ---------------------------------------------------------- 1040

Db    1710 TGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATT 1769

Qy    1041 ---------------------------------------------------------- 1040

Db    1770 TTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTG 1829

Qy    1041 ---------------------------------------------------------- 1040

Db    1830 TATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG 1889

Qy    1041 ---------------------------------------------------------- 1040

Db    1890 AAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGC 1949

Qy    1041 ----------------------------------tgcggtgggctctatggcttct-- 1062
                                               |||| || ||||  || |  ||
Db    1950 CTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTT 2009

Qy    1063 --------gaggcggaaagaaccagctgcattaatgaatcggccaacgcgcgggggagagg 1114
                   ||  |      |  ||||||||||||||||||||||||||||||||||||||
Db    2010 CCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGG 2069

Qy    1115 cggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgt 1174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2070 CGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGT 2129

Qy    1175 tcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatc 1234
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2130 TCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC 2189

Qy    1235 aggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaa 1294
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2190 AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAA 2249

Qy    1295 aaaggccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaa 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2250 AAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCCTGACGAGCATCACAAAAA 2309

Qy    1355 tcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcc 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2310 TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCC 2369

Qy    1415 ccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtc 1474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2370 CCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC 2429

Qy    1475 cgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcag 1534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    2430 CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTGAG 2489

Qy    1535 ttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccga 1594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2490 TTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGA 2549

Qy    1595 ccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatc 1654
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2550 CCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC 2609

Qy    1655 gccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctac 1714
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2610 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC 2669

Qy    1715 agagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctg 1774
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2670 AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTG 2729

Qy    1775 cgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaaca 1834
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2730 CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACA 2789

Qy    1835 aaccaccgctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaa 1894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2790 AACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAA 2849

Qy    1895 aggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa 1954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2850 AGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAA 2909

Qy    1955 ctcacgttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          |||||||||||||||||||||||||   ||  | |      |  |  ||||
Db    2910 CTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 2962


RESULT   3
US-08-760-615-7
; Sequence 7, Application US/08760615
; Patent No. 6200959
;  GENERAL INFORMATION:
;    APPLICANT:  Haynes, Joel R
;    APPLICANT:  Schmaljohn, Connie S
;    APPLICANT:  Fuller, Deborah L
;    APPLICANT:  Schmaljohn, Alan
;    APPLICANT:  Jahrling, Peter B
;    TITLE OF INVENTION:  GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
;    TITLE OF INVENTION:  RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53703
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/760,615
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berson, Bennett J
;      REGISTRATION NUMBER:  37094
;      REFERENCE/DOCKET NUMBER:  110229.91241
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4326 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Expression vector
;      DESCRIPTION:  construct"
;    IMMEDIATE SOURCE:
;      CLONE:  pWRG7077
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  1250..2062
;    FEATURE:
;      NAME/KEY:  intron
;      LOCATION:  2063..2887
;      OTHER INFORMATION:  /function= "Human Cytomegalovirus
;      OTHER INFORMATION:  Intron A"
;    FEATURE:
;      NAME/KEY:  polyA_site
;      LOCATION:  2912..3314
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (299..1114)
US-08-760-615-7


  Query Match             13.0%;  Score 1053.6;  DB 4;  Length 4326;
  Best Local Similarity   94.1%;  Pred. No. 7.6e-227;
  Matches 1118;  Conservative    0;  Mismatches   39;  Indels   31;  Gaps    1;

Qy     820 tgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtg 879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3030 TGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTG 3089

Qy     880 ccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaatt 939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3090 CCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATT 3149

Qy     940 gcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagc 999
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3150 GCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGC 3209

Qy    1000 aaggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggct 1059
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3210 AAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCT 3269

Qy    1060 tctgaggcggaaagaaccagctgcattaatgaatcggccaacgcgcggggagaggcggtt 1119
           ||||||||||||||||||||                                || | |
Db    3270 TCTGAGGCGGAAAGAACCAGCT------------------------------GGGGCTC 3298

Qy    1120 tgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggc 1179
             |   | |           ||||||||||||||||||||||||||||||||||||||||||||
Db    3299 GACAGCTCGACTCTAGAATTGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGC 3358

Qy    1180 tgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcagggg 1239
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3359 TGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGG 3418

Qy   1240 ataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaagg 1299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3419 ATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGG 3478

Qy   1300 ccgcgttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgac 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3479 CCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC 3538

Qy   1360 gctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctg 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3539 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTG 3598

Qy   1420 gaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcct 1479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3599 GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCT 3658

Qy   1480 ttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcgg 1539
          |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db   3659 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGG 3718

Qy   1540 tgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgct 1599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3719 TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT 3778

Qy   1600 gcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac 1659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3779 GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC 3838

Qy   1660 tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagt 1719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3839 TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT 3898

Qy   1720 tcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctc 1779
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   3899 TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTC 3958

Qy   1780 tgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaacca 1839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3959 TGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCA 4018

Qy   1840 ccgctggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggat 1899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4019 CCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGAT 4078

Qy   1900 ctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcac 1959
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4079 CTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC 4138

Qy   1960 gttaagggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          |||||||||||||||||||  ||  | |      |  |  ||||
Db   4139 GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 4186


RESULT   4
US-08-893-327-15
; Sequence 15, Application US/08893327
; Patent No. 6020192
;  GENERAL INFORMATION:
;    APPLICANT:  Zolotukhin, Sergei
;    APPLICANT:  Hauswirth, William W.
;    APPLICANT:  Muzyczka, Nicholas
;    TITLE OF INVENTION:  Humanized Green Fluorescent Protein
;    TITLE OF INVENTION:  Genes and Methods
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P. O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/893,327
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/588,201
;      FILING DATE:  18-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  UFLA:062\KIT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6253 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  988..1701
US-08-893-327-15


  Query Match             12.1%;  Score 981.8;  DB 3;  Length 6253;
  Best Local Similarity   86.9%;  Pred. No. 1.1e-210;
  Matches 1162;  Conservative    0;  Mismatches   37;  Indels  138;  Gaps    2;

Qy    809 ctagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcc 868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3017 CTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC 3076

Qy    869 cctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaa 928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3077 CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA 3136

Qy    929 atgaggaaattgcatcgcattgtctgagtaggtgtcattctattctggggggtggggtgg 988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3137 ATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG 3196

Qy    989 ggcaggacagcaaggggggaggattgggaagacaatagcaggcatgctgggga-------- 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||--------
Db   3197 GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAGAGATCTA 3256

Qy   1041 ------------------------------------------------------------ 1040

Db   3257 GGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGC 3316

Qy   1041 ------------------tgcggtgggctctatggcttctgaggcggaaagaaccag-- 1079
                             |  | | |  || |||   |   | |  |  || | ||
Db   3317 CGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCG 3376

Qy   1080 ----------------------------------------------------ctgcattaatg 1090
                                                             ||||||||||||
Db   3377 AGCGCGCAGAGAGGGAGTGGCCAACCCCCCCCCCCCCCCCCCTGCAGCCCTGCATTAATG 3436

Qy   1091 aatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgct 1150
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3437 AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT 3496

Qy   1151 cactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggc 1210
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3497 CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC 3556

Qy   1211 ggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaagg 1270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3557 GGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGG 3616

Qy    1271 ccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccg 1330
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3617 CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG 3676

Qy    1331 ccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagg 1390
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3677 CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG 3736

Qy    1391 actataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgac 1450
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3737 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC 3796

Qy    1451 cctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctca 1510
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3797 CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA 3856

Qy    1511 atgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgt 1570
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3857 ATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGT 3916

Qy    1571 gcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtc 1630
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3917 GCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC 3976

Qy    1631 caacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag 1690
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3977 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG 4036

Qy    1691 agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacac 1750
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4037 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC 4096

Qy    1751 tagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagt 1810
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4097 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGT 4156

Qy    1811 tggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaa 1870
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4157 TGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAA 4216

Qy    1871 gcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggg 1930
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4217 GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG 4276

Qy    1931 gtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacat 1990
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4277 GTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA 4336

Qy    1991 atttgaatgtatttaga 2007
           |      |  |  ||||
Db    4337 AAGGATCTTCACCTAGA 4353


RESULT   5
US-08-893-327-17
; Sequence 17, Application US/08893327
; Patent No. 6020192
;  GENERAL INFORMATION:
;    APPLICANT:  Zolotukhin, Sergei
;    APPLICANT:  Hauswirth, William W.
;    APPLICANT:  Muzyczka, Nicholas
;    TITLE OF INVENTION:  Humanized Green Fluorescent Protein
;    TITLE OF INVENTION:  Genes and Methods
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P. O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/893,327
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/588,201
;      FILING DATE:  18-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  UFLA:062\KIT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6280 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  988..1728
US-08-893-327-17


  Query Match             12.1%;  Score 981.8;  DB 3;  Length 6280;
  Best Local Similarity   86.9%;  Pred. No. 1.1e-210;
  Matches 1162;  Conservative    0;  Mismatches   37;  Indels  138;  Gaps    2;

Qy     809 ctagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcc 868
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3044 CTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC 3103

Qy     869 cctcccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaa 928
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3104 CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA 3163

Qy     929 atgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtgg 988
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3164 ATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG 3223

Qy     989 ggcaggacagcaaggggaggattgggaagacaatagcaggcatgctgggga-------- 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3224 GGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAGAGATCTA 3283

Qy    1041 ---------------------------------------------------------- 1040

Db    3284 GGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGC 3343

Qy    1041 --------------------tgcggtgggctctatggcttctgaggcggaaagaaccag-- 1079
                               | | | | || |||   |   | | |  || | ||
Db    3344 CGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCG 3403

Qy    1080 ----------------------------------------------------ctgcattaatg 1090
                                                               ||||||||||||
Db    3404 AGCGCGCAGAGAGGGAGTGGCCAACCCCCCCCCCCCCCCCCTGCAGCCCTGCATTAATG 3463

Qy    1091 aatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgct 1150
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3464 AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT 3523

Qy    1151 cactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggc 1210
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3524 CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC 3583

Qy    1211 ggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaagg 1270
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3584 GGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGG 3643

Qy   1271 ccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccg 1330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3644 CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG 3703

Qy   1331 ccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagg 1390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3704 CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG 3763

Qy   1391 actataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgac 1450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3764 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC 3823

Qy   1451 cctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctca 1510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3824 CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA 3883

Qy   1511 atgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgt 1570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3884 ATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGT 3943

Qy   1571 gcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtc 1630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3944 GCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC 4003

Qy   1631 caacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag 1690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4004 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG 4063

Qy   1691 agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacac 1750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4064 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC 4123

Qy   1751 tagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagt 1810
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4124 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGT 4183

Qy   1811 tggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaa 1870
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4184 TGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTGTTTGCAA 4243

Qy   1871 gcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggg 1930
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4244 GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG 4303

Qy   1931 gtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacat 1990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||  |
Db   4304 GTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA 4363

Qy   1991 atttgaatgtatttaga 2007
          |      |  |  ||||
Db   4364 AAGGATCTTCACCTAGA 4380


RESULT   6
US-08-893-327-19
; Sequence 19, Application US/08893327
; Patent No. 6020192
;  GENERAL INFORMATION:
;    APPLICANT:  Zolotukhin, Sergei
;    APPLICANT:  Hauswirth, William W.
;    APPLICANT:  Muzyczka, Nicholas
;    TITLE OF INVENTION:  Humanized Green Fluorescent Protein
;    TITLE OF INVENTION:  Genes and Methods
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P. O. Box 4433
;      CITY:  Houston
;      STATE:  TX
;      COUNTRY:  USA
;      ZIP:  77210-4433
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/893,327
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/588,201
;      FILING DATE:  18-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kitchell, Barbara S.
;      REGISTRATION NUMBER:  33,928
;      REFERENCE/DOCKET NUMBER:  UFLA:062\KIT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6280 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  988..1728
US-08-893-327-19


  Query Match             12.1%;  Score 981.8;  DB 3;  Length 6280;
  Best Local Similarity   86.9%;  Pred. No. 1.1e-210;
  Matches 1162;  Conservative    0;  Mismatches   37;  Indels  138;  Gaps    2;

Qy    809 ctagagctcgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcc 868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3044 CTAGAGCTCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCC 3103

Qy    869 cctcccccgtgccttccttgaccctggaaggtgccactcccactgtccttttcctaataaa 928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3104 CCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAA 3163

Qy    929 atgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtgg 988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3164 ATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGG 3223

Qy    989 ggcaggacagcaaggggaggattgggaagacaatagcaggcatgctgggga-------- 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||        
Db   3224 GGCAGGACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAGAGATCTA 3283

Qy   1041 ---------------------------------------------------------- 1040

Db   3284 GGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGC 3343

Qy   1041 ------------------tgcggtgggctctatggcttctgaggcggaaagaaccag-- 1079
                            |  | | |  || |||   |   | |  |  || | ||  
Db   3344 CGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCG 3403

Qy   1080 ----------------------------------------------------ctgcattaatg 1090
                                                               ||||||||||| 
Db   3404 AGCGCGCAGAGAGGGAGTGGCCAACCCCCCCCCCCCCCCCCCTGCAGCCCTGCATTAATG 3463

Qy   1091 aatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgct 1150
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3464 AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT 3523

Qy   1151 cactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggc 1210
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3524 CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC 3583

Qy   1211 ggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaagg 1270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3584 GGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGG 3643

Qy    1271 ccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccg 1330
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3644 CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG 3703

Qy    1331 ccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagg 1390
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3704 CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG 3763

Qy    1391 actataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgac 1450
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3764 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC 3823

Qy    1451 cctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctca 1510
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3824 CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA 3883

Qy    1511 atgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgt 1570
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3884 ATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGT 3943

Qy    1571 gcacgaacccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtc 1630
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3944 GCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC 4003

Qy    1631 caacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag 1690
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4004 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG 4063

Qy    1691 agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacac 1750
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4064 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC 4123

Qy    1751 tagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagt 1810
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4124 TAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGT 4183

Qy    1811 tggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaa 1870
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4184 TGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAA 4243

Qy    1871 gcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggg 1930
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4244 GCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG 4303

Qy    1931 gtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacat 1990
           |||||||||||||||||||||||||||||||||||||||||||||||||||||  ||  |
Db    4304 GTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA 4363

Qy    1991 atttgaatgtatttaga 2007
           |      |  |  ||||
Db    4364 AAGGATCTTCACCTAGA 4380


RESULT   7
US-08-343-401A-3
; Sequence 3, Application US/08343401A
; Patent No. 5661132
;  GENERAL INFORMATION:
;    APPLICANT:  Swain, William F
;    APPLICANT:  Macklin, Michael D
;    APPLICANT:  Eriksson, Elof
;    APPLICANT:  Andree, Christophe
;    TITLE OF INVENTION:  Improved Wound Healing
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  PO Box 2113
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  USA
;      ZIP:  53701-2113
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/343,401A
;      FILING DATE:  22-NOV-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J
;      REGISTRATION NUMBER:  27,386
;      REFERENCE/DOCKET NUMBER:  11-229-9103-9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4283 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE:  pWRG1630
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  713..721
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  981..1253
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(713..721, 981..1253)
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  713..1049
US-08-343-401A-3


  Query Match             12.0%;  Score 973;  DB 1;  Length 4283;
  Best Local Similarity   83.2%;  Pred. No. 8.6e-209;
  Matches 1216;  Conservative    0;  Mismatches   30;  Indels  215;  Gaps    1;

Qy     762 ctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctg 821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1277 CTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTG 1336

Qy     822 atcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgcc 881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1337 ATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCC 1396

Qy     882 ttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgc 941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1397 TTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC 1456

Qy     942 atcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaa 1001
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1457 ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA 1516

Qy    1002 gggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttc 1061
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db    1517 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAACC 1576

Qy    1062 tgaggcggaaaga---------------------------------------------- 1074
           |  | ||   ||
Db    1577 AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT 1636

Qy    1075 ---------------------------------------------------------- 1074

Db    1637 AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA 1696

Qy    1075 ---------------------------------------------------------- 1074
```

```
Db   1697 TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT 1756

Qy   1075 ----------------------------------------------accagctgcatt 1086
                                                        ||||||||||||
Db   1757 TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATT 1816

Qy   1087 aatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcct 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1817 AATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT 1876

Qy   1147 cgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaa 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1877 CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 1936

Qy   1207 aggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaa 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1937 AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA 1996

Qy   1267 aaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggc 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1997 AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCGATAGGC 2056

Qy   1327 tccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccga 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2057 TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 2116

Qy   1387 caggactataaagataccaggcgtttcccctggaagctccctcgtgcgctctcctgttc 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117 CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC 2176

Qy   1447 cgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgcttt 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177 CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 2236

Qy   1507 ctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct 1566
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237 CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT 2296

Qy   1567 gtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297 GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG 2356

Qy   1627 agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatta 1686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA 2416

Qy   1687 gcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggct 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2417 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCT 2476

Qy   1747 acactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2477 ACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 2536

Qy   1807 gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgttt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2537 GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT 2596

Qy   1867 gcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttcta 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2597 GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA 2656

Qy   1927 cggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggat 1986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db   2657 CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 2716

Qy   1987 acatatttgaatgtatttaga 2007
            | |     |  |  ||||
Db   2717 CAAAAAGGATCTTCACCTAGA 2737

RESULT   8
```

```
US-08-445-265A-1
; Sequence 1, Application US/08445265A
; Patent No. 5697901
;  GENERAL INFORMATION:
;    APPLICANT:  Eriksson, Elof
;    TITLE OF INVENTION:  GENE DELIVERY BY MICRONEEDLE INJECTION
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53703
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/445,265A
;      FILING DATE:
;      CLASSIFICATION:  604
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seay, Nicholas J
;      REGISTRATION NUMBER:  27386
;      REFERENCE/DOCKET NUMBER:  110229.91080
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4283 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Plasmid DNA"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(713..721, 981..1250)
US-08-445-265A-1


  Query Match             12.0%;  Score 973;  DB 1;  Length 4283;
  Best Local Similarity   83.2%;  Pred. No. 8.6e-209;
  Matches 1216;  Conservative    0;  Mismatches   30;  Indels  215;  Gaps    1;

Qy    762 ctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctg 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1277 CTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTG 1336

Qy    822 atcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtgcc 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1337 ATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCC 1396

Qy    882 ttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgc 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1397 TTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC 1456

Qy    942 atcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaa 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1457 ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA 1516

Qy   1002 gggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttc 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db   1517 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAACC 1576

Qy   1062 tgaggcggaaaga---------------------------------------------- 1074
           |  | ||   ||
Db   1577 AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT 1636

Qy   1075 ---------------------------------------------------------- 1074
```

```
Db   1637 AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA 1696

Qy   1075 ------------------------------------------------------------ 1074

Db   1697 TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT 1756

Qy   1075 ----------------------------------------------accagctgcatt 1086
                                                          ||||||||||||
Db   1757 TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATT 1816

Qy   1087 aatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcct 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1817 AATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT 1876

Qy   1147 cgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaa 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1877 CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 1936

Qy   1207 aggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaa 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1937 AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA 1996

Qy   1267 aaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggc 1326
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1997 AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCGATAGGC 2056

Qy   1327 tccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccga 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2057 TCCGCCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 2116

Qy   1387 caggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttc 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117 CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC 2176

Qy   1447 cgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgcttt 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177 CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 2236

Qy   1507 ctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct 1566
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237 CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT 2296

Qy   1567 gtgtgcacgaacccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297 GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG 2356

Qy   1627 agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatta 1686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA 2416

Qy   1687 gcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggct 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2417 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCT 2476

Qy   1747 acactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2477 ACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 2536

Qy   1807 gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgttt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2537 GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT 2596

Qy   1867 gcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttcta 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2597 GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA 2656

Qy   1927 cggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggat 1986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db   2657 CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 2716

Qy   1987 acatatttgaatgtatttaga 2007
            | |      |  |  ||||
```

```
Db   2717 CAAAAAGGATCTTCACCTAGA 2737


RESULT   9
US-08-990-442-1
; Sequence 1, Application US/08990442
; Patent No. 6090790
;  GENERAL INFORMATION:
;    APPLICANT:  Eriksson, Elof
;    TITLE OF INVENTION:  GENE DELIVERY BY MICRONEEDLE INJECTION
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  1 South Pinckney Street
;      CITY:  Madison
;      STATE:  WI
;      COUNTRY:  US
;      ZIP:  53703
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/990,442
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berson, Bennett J
;      REGISTRATION NUMBER:  37094
;      REFERENCE/DOCKET NUMBER:  310558.90028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  608-251-5000
;      TELEFAX:  608-251-9166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4283 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Plasmid DNA"
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(713..721, 981..1250)
US-08-990-442-1


  Query Match             12.0%;  Score 973;  DB 3;  Length 4283;
  Best Local Similarity   83.2%;  Pred. No. 8.6e-209;
  Matches 1216;  Conservative    0;  Mismatches   30;  Indels  215;  Gaps    1;

Qy    762 ctcgagcatgcatctagagggccctattctatagtgtcacctaaatgctagagctcgctg 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1277 CTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTG 1336

Qy    822 atcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgcc 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1337 ATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCC 1396

Qy    882 ttccttgaccctggaaggtgccactcccactgtccttcctaataaaatgaggaaattgc 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1397 TTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGC 1456

Qy    942 atcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaa 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1457 ATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA 1516

Qy   1002 gggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttc 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||    |
Db   1517 GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAACC 1576

Qy   1062 tgaggcggaaaga---------------------------------------------- 1074
```

```
          |  | ||   ||
Db   1577 AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT 1636

Qy   1075 ------------------------------------------------------------ 1074

Db   1637 AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA 1696

Qy   1075 ------------------------------------------------------------ 1074

Db   1697 TACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT 1756

Qy   1075 ----------------------------------------------accagctgcatt 1086
                                                         ||||||||||||
Db   1757 TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATT 1816

Qy   1087 aatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcct 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1817 AATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT 1876

Qy   1147 cgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaa 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1877 CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA 1936

Qy   1207 aggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaa 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1937 AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA 1996

Qy   1267 aaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggc 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   1997 AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCGATAGGC 2056

Qy   1327 tccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccga 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2057 TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGA 2116

Qy   1387 caggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttc 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2117 CAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC 2176

Qy   1447 cgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgcttt 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2177 CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT 2236

Qy   1507 ctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct 1566
          ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2237 CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT 2296

Qy   1567 gtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttg 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2297 GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG 2356

Qy   1627 agtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatta 1686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2357 AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTA 2416

Qy   1687 gcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggct 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2417 GCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCT 2476

Qy   1747 acactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaa 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2477 ACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA 2536

Qy   1807 gagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgttt 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2537 GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTT 2596

Qy   1867 gcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttcta 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2597 GCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA 2656

Qy   1927 cggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggat 1986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db   2657 CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT 2716

Qy   1987 acatatttgaatgtatttaga 2007
          | ||      |  |  ||||
Db   2717 CAAAAAGGATCTTCACCTAGA 2737
```

RESULT 10
US-08-467-420A-49
; Sequence 49, Application US/08467420A
; Patent No. 5683892
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert S.
; APPLICANT: Appelbaum, Edward R.
; APPLICANT: Chaiken, Irwin M.
; APPLICANT: Cook, Richard M.
; APPLICANT: Gross, Mitchell S.
; APPLICANT: Holmes, Stephen D.
; APPLICANT: McMillan, Lynette J.
; APPLICANT: Theisen, Timothy W.
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
; TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; ADDRESSEE: Intellectual Property
; STREET: P. O. Box 1539-UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,420A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363131
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P50282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610 270-5024
; TELEFAX: 610 270-5090
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-467-420A-49

Query Match 11.4%; Score 921.8; DB 1; Length 6285;
Best Local Similarity 82.9%; Pred. No. 2.9e-197;
Matches 1160; Conservative 0; Mismatches 27; Indels 212; Gaps 1;

```
Qy    821 gatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgc 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3829 GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC 3888

Qy    881 cttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3889 CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG 3948

Qy    941 catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca 1000
```

```
Db   3949 CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCA 4008

Qy   1001 agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggctt 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAC 4068

Qy   1061 ctgaggcggaaaga---------------------------------------------- 1074
          | |  | ||   ||
Db   4069 CAGCTGGGGCTCGATCGAGTGTATGACTGCGGCCGCGATCCCGTCGAGAGCTTGGCGTAA 4128

Qy   1075 ---------------------------------------------------------- 1074

Db   4129 TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA 4188

Qy   1075 ---------------------------------------------------------- 1074

Db   4189 CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA 4248

Qy   1075 ----------------------------------------------accagctgcattaa 1088
                                                          ||||||||||||||
Db   4249 ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA 4308

Qy   1089 tgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg 1148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG 4368

Qy   1149 ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag 1208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4369 CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 4428

Qy   1209 gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaa 1268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4429 GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA 4488

Qy   1269 ggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc 1328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4489 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC 4548

Qy   1329 cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca 1388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4549 CGCCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 4608

Qy   1389 ggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccg 1448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4609 GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG 4668

Qy   1449 accctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct 1508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT 4728

Qy   1509 caatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt 1568
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT 4788

Qy   1569 gtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag 1628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4789 GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG 4848

Qy   1629 tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc 1688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4849 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC 4908

Qy   1689 agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac 1748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4909 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC 4968

Qy   1749 actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga 1808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4969 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5028

Qy   1809 gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgc 1868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   5029 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC 5088

Qy   1869 aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg 1928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5089 AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG 5148

Qy   1929 gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatac 1988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db   5149 GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCA 5208

Qy   1989 atatttgaatgtatttaga 2007
          | |      |  |  ||||
Db   5209 AAAAGGATCTTCACCTAGA 5227


RESULT 11
US-08-470-110A-49
; Sequence 49, Application US/08470110A
; Patent No. 5693323
;  GENERAL INFORMATION:
;    APPLICANT:  Ames, Robert S.
;    APPLICANT:  Appelbaum, Edward R.
;    APPLICANT:  Chaiken, Irwin M.
;    APPLICANT:  Cook, Richard M.
;    APPLICANT:  Gross, Mitchell S.
;    APPLICANT:  Holmes, Stephen D.
;    APPLICANT:  McMillan, Lynette J.
;    APPLICANT:  Theisen, Timothy W.
;    TITLE OF INVENTION:  Recombinant IL5 Antagonists Useful in
;    TITLE OF INVENTION:  Treatment of IL5 Mediated Disorders
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      ADDRESSEE:  Intellectual Property
;      STREET:  P. O. Box 1539-UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,110A
;      FILING DATE:
;      CLASSIFICATION:  426
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/363131
;      FILING DATE:  23-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sutton, Jeffrey A.
;      REGISTRATION NUMBER:  34,028
;      REFERENCE/DOCKET NUMBER:  P50282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610 270-5024
;      TELEFAX:  610 270-5090
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-470-110A-49


  Query Match             11.4%;  Score 921.8;  DB 1;  Length 6285;
  Best Local Similarity   82.9%;  Pred. No. 2.9e-197;
  Matches 1160;  Conservative    0;  Mismatches   27;  Indels  212;  Gaps    1;

Qy    821 gatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtgc 880
```

```
Db   3829 GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC 3888

Qy    881 cttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg 940
Db   3889 CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG 3948

Qy    941 catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca 1000
Db   3949 CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCA 4008

Qy   1001 agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggctt 1060
Db   4009 AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAC 4068

Qy   1061 ctgaggcggaaaga---------------------------------------------- 1074
Db   4069 CAGCTGGGGCTCGATCGAGTGTATGACTGCGGCCGCGATCCCGTCGAGAGCTTGGCGTAA 4128

Qy   1075 ---------------------------------------------------------- 1074

Db   4129 TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA 4188

Qy   1075 ---------------------------------------------------------- 1074

Db   4189 CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA 4248

Qy   1075 ----------------------------------------------accagctgcattaa 1088
Db   4249 ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA 4308

Qy   1089 tgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg 1148
Db   4309 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG 4368

Qy   1149 ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag 1208
Db   4369 CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 4428

Qy   1209 gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaa 1268
Db   4429 GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA 4488

Qy   1269 ggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc 1328
Db   4489 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC 4548

Qy   1329 cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca 1388
Db   4549 CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 4608

Qy   1389 ggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccg 1448
Db   4609 GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG 4668

Qy   1449 accctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct 1508
Db   4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT 4728

Qy   1509 caatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt 1568
Db   4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT 4788

Qy   1569 gtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag 1628
Db   4789 GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG 4848

Qy   1629 tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc 1688
Db   4849 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC 4908

Qy   1689 agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac 1748
Db   4909 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC 4968

Qy   1749 actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga 1808
Db   4969 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5028

Qy   1809 gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgc 1868
Db   5029 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC 5088

Qy   1869 aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg 1928
Db   5089 AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG 5148

Qy   1929 gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatac 1988
Db   5149 GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCA 5208

Qy   1989 atatttgaatgtatttaga 2007
Db   5209 AAAAGGATCTTCACCTAGA 5227
```

RESULT 12
US-08-667-769A-49
; Sequence 49, Application US/08667769A
; Patent No. 5783184
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert S.
; APPLICANT: Appelbaum, Edward R.
; APPLICANT: Chaiken, Irwin M.
; APPLICANT: Cook, Richard M.
; APPLICANT: Gross, Mitchell S.
; APPLICANT: Holmes, Stephen D.
; APPLICANT: McMillan, Lynette J.
; APPLICANT: Theisen, Timothy W.
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
; TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; STREET: P.O. Box 1539-UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/667,769A
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17082
; FILING DATE: 22-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470110
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/467420
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363131
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P50503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5024

```
;      TELEFAX:  610-270-5090
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-667-769A-49


  Query Match           11.4%;  Score 921.8;  DB 1;  Length 6285;
  Best Local Similarity  82.9%;  Pred. No. 2.9e-197;
  Matches 1160;  Conservative    0;  Mismatches   27;  Indels  212;  Gaps    1;

Qy    821 gatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgc 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3829 GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC 3888

Qy    881 cttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3889 CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG 3948

Qy    941 catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3949 CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCA 4008

Qy   1001 agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggctt 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAC 4068

Qy   1061 ctgaggcggaaaga---------------------------------------------- 1074
          | | | ||   ||
Db   4069 CAGCTGGGGCTCGATCGAGTGTATGACTGCGGCCGCGATCCCGTCGAGAGCTTGGCGTAA 4128

Qy   1075 ---------------------------------------------------------- 1074

Db   4129 TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA 4188

Qy   1075 ---------------------------------------------------------- 1074

Db   4189 CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA 4248

Qy   1075 ----------------------------------------------accagctgcattaa 1088
                                                       |||||||||||||
Db   4249 ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA 4308

Qy   1089 tgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg 1148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG 4368

Qy   1149 ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag 1208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4369 CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 4428

Qy   1209 gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaa 1268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4429 GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA 4488

Qy   1269 ggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc 1328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4489 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC 4548

Qy   1329 cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca 1388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4549 CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 4608

Qy   1389 ggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccg 1448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4609 GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG 4668

Qy   1449 accctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct 1508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT 4728

Qy   1509 caatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt 1568
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT 4788

Qy   1569 gtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag 1628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4789 GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG 4848

Qy   1629 tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc 1688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4849 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC 4908

Qy   1689 agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac 1748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4909 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC 4968

Qy   1749 actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga 1808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4969 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5028

Qy   1809 gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttttgtttgc 1868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5029 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGC 5088

Qy   1869 aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg 1928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5089 AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG 5148

Qy   1929 gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatac 1988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db   5149 GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCA 5208

Qy   1989 atatttgaatgtatttaga 2007
          | |       |  |  ||||
Db   5209 AAAAGGATCTTCACCTAGA 5227


RESULT  13
US-08-940-371-49
; Sequence 49, Application US/08940371
; Patent No. 5851525
;  GENERAL INFORMATION:
;    APPLICANT:  Ames, Robert S.
;    APPLICANT:  Appelbaum, Edward R.
;    APPLICANT:  Chaiken, Irwin M.
;    APPLICANT:  Cook, Richard M.
;    APPLICANT:  Gross, Mitchell S.
;    APPLICANT:  Holmes, Stephen D.
;    APPLICANT:  McMillan, Lynette J.
;    APPLICANT:  Theisen, Timothy W.
;    TITLE OF INVENTION:  Recombinant IL5 Antagonists Useful in
;    TITLE OF INVENTION:  Treatment of IL5 Mediated Disorders
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      ADDRESSEE:  Intellectual Property
;      STREET:  P. O. Box 1539-UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,371
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/470,110
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/363131
;      FILING DATE:  23-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sutton, Jeffrey A.
;      REGISTRATION NUMBER:  34,028
;      REFERENCE/DOCKET NUMBER:  P50282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610 270-5024
;      TELEFAX:  610 270-5090
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
US-08-940-371-49


  Query Match             11.4%;  Score 921.8;  DB 2;  Length 6285;
  Best Local Similarity   82.9%;  Pred. No. 2.9e-197;
  Matches 1160;  Conservative    0;  Mismatches   27;  Indels  212;  Gaps    1;

Qy    821 gatcagcctcgactgtgccttctagttgccagccatctgttgtttgccccctcccccgtgc 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3829 GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC 3888

Qy    881 cttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3889 CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG 3948

Qy    941 catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3949 CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCA 4008

Qy   1001 agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggctt 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4009 AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAC 4068

Qy   1061 ctgaggcggaaaga---------------------------------------------- 1074
          | |  | ||   ||
Db   4069 CAGCTGGGGCTCGATCGAGTGTATGACTGCGGCCGCGATCCCGTCGAGAGCTTGGCGTAA 4128

Qy   1075 ---------------------------------------------------------- 1074

Db   4129 TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA 4188

Qy   1075 ---------------------------------------------------------- 1074

Db   4189 CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA 4248

Qy   1075 ----------------------------------------------accagctgcattaa 1088
                                                         ||||||||||||||
Db   4249 ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA 4308

Qy   1089 tgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg 1148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4309 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG 4368

Qy   1149 ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag 1208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4369 CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 4428

Qy   1209 gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaa 1268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4429 GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA 4488

Qy   1269 ggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc 1328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4489 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC 4548

Qy   1329 cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca 1388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4549 CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 4608

Qy   1389 ggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccg 1448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4609 GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG 4668

Qy   1449 accctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct 1508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT 4728

Qy   1509 caatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt 1568
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT 4788

Qy   1569 gtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag 1628
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4789 GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG 4848

Qy   1629 tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc 1688
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4849 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC 4908

Qy   1689 agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac 1748
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4909 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC 4968

Qy   1749 actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga 1808
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4969 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5028

Qy   1809 gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgc 1868
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5029 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC 5088

Qy   1869 aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg 1928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5089 AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG 5148

Qy   1929 gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatac 1988
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db   5149 GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCA 5208

Qy   1989 atatttgaatgtatttaga 2007
          | |      |  |  ||||
Db   5209 AAAAGGATCTTCACCTAGA 5227


RESULT  14
US-08-637-647-49
; Sequence 49, Application US/08637647
; Patent No. 6129913
;  GENERAL INFORMATION:
;    APPLICANT:  Ames, Robert S.
;    APPLICANT:  Appelbaum, Edward R.
;    APPLICANT:  Chaiken, Irwin M.
;    APPLICANT:  Cook, Richard M.
;    APPLICANT:  Gross, Mitchell S.
;    APPLICANT:  McMillan, Lynette J.
;    APPLICANT:  Theisen, Timothy W.
;    TITLE OF INVENTION:  Recombinant IL5 Antagonists Useful in
;    TITLE OF INVENTION:  Treatment of IL5 Mediated Disorders
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      ADDRESSEE:  Intellectual Property
;      STREET:  P. O. Box 1539-UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/637,647
;       FILING DATE:  19-JUN-1997
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/363,131
;       FILING DATE:
;       CLASSIFICATION:  536
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sutton, Jeffrey A.
;       REGISTRATION NUMBER:  34,028
;       REFERENCE/DOCKET NUMBER:  P50282
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  610 270-5024
;       TELEFAX:  610 270-5090
;   INFORMATION FOR SEQ ID NO:  49:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6285 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  circular
;     MOLECULE TYPE:  DNA (genomic)
US-08-637-647-49


  Query Match             11.4%;  Score 921.8;  DB 3;  Length 6285;
  Best Local Similarity   82.9%;  Pred. No. 2.9e-197;
  Matches 1160;  Conservative    0;  Mismatches   27;  Indels  212;  Gaps    1;

Qy.    821 gatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccccgtgc 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3829 GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCCGTGC 3888

Qy     881 cttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3889 CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG 3948

Qy     941 catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3949 CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCA 4008

Qy    1001 aggggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggctt 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4009 AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAC 4068

Qy    1061 ctgaggcggaaaga---------------------------------------------- 1074
           | |  | ||   ||
Db    4069 CAGCTGGGGCTCGATCGAGTGTATGACTGCGGCCGCGATCCCGTCGAGAGCTTGGCGTAA 4128

Qy    1075 ---------------------------------------------------------- 1074

Db    4129 TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA 4188

Qy    1075 ---------------------------------------------------------- 1074

Db    4189 CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA 4248

Qy    1075 ----------------------------------------------accagctgcattaa 1088
                                                           ||||||||||||||
Db    4249 ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA 4308

Qy    1089 tgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg 1148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4309 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG 4368

Qy    1149 ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag 1208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4369 CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 4428

Qy    1209 gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaa 1268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4429 GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA 4488

Qy    1269 ggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc 1328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4489 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC 4548

Qy    1329 cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca 1388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4549 CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 4608

Qy    1389 ggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccg 1448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4609 GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG 4668

Qy    1449 accctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct 1508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT 4728

Qy    1509 caatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt 1568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT 4788

Qy    1569 gtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag 1628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4789 GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG 4848

Qy    1629 tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc 1688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4849 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC 4908

Qy    1689 agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac 1748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4909 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC 4968

Qy    1749 actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga 1808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4969 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5028

Qy    1809 gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgc 1868
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5029 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC 5088

Qy    1869 aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg 1928
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5089 AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG 5148

Qy    1929 gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatac 1988
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db    5149 GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCA 5208

Qy    1989 atatttgaatgtatttaga 2007
           | |       |  |  ||||
Db    5209 AAAAGGATCTTCACCTAGA 5227


RESULT  15
PCT-US95-17082A-49
; Sequence 49, Application PC/TUS9517082A
;  GENERAL INFORMATION:
;     APPLICANT:  Ames, Robert S.
;     APPLICANT:  Appelbaum, Edward R.
;     APPLICANT:  Chaiken, Irwin M.
;     APPLICANT:  Cook, Richard M.
;     APPLICANT:  Gross, Mitchell S.
;     APPLICANT:  Holmes, Stephen D.
;     APPLICANT:  McMillan, Lynette J.
;     APPLICANT:  Theisen, Timothy W.
;     TITLE OF INVENTION:  Recombinant IL5 Antagonists Useful in
;     TITLE OF INVENTION:  Treatment of IL5 Mediated Disorders
;     NUMBER OF SEQUENCES:  76
;     CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      STREET:  P.O. Box 1539-UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17082A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/470110
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/467420
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/363131
;      FILING DATE:  23-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sutton, Jeffrey A.
;      REGISTRATION NUMBER:  34,028
;      REFERENCE/DOCKET NUMBER:  P50282-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5024
;      TELEFAX:  610-270-5090
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6285 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
PCT-US95-17082A-49


  Query Match             11.4%;  Score 921.8;  DB 5;  Length 6285;
  Best Local Similarity   82.9%;  Pred. No. 2.9e-197;
  Matches 1160;  Conservative    0;  Mismatches   27;  Indels  212;  Gaps    1;

Qy     821 gatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtgc 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3829 GATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGC 3888

Qy     881 cttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattg 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3889 CTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTG 3948

Qy     941 catcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagca 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3949 CATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCA 4008

Qy    1001 agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggctt 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4009 AGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGAAC 4068

Qy    1061 ctgaggcggaaaga---------------------------------------------- 1074
           | |  | ||   ||
Db    4069 CAGCTGGGGCTCGATCGAGTGTATGACTGCGGCCGCGATCCCGTCGAGAGCTTGGCGTAA 4128

Qy    1075 ---------------------------------------------------------- 1074

Db    4129 TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA 4188

Qy    1075 ---------------------------------------------------------- 1074

Db    4189 CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA 4248

Qy    1075 ----------------------------------------------accagctgcattaa 1088
                                                           ||||||||||||
Db    4249 ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA 4308

Qy    1089 tgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcg 1148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4309 TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCG 4368

Qy   -1149 ctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaag 1208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4369 CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG 4428

Qy    1209 gcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaa 1268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4429 GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAA 4488

Qy    1269 ggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctc 1328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4489 GGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC 4548

Qy    1329 cgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgaca 1388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4549 CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 4608

Qy    1389 ggactataaagatacccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccg 1448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4609 GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG 4668

Qy    1449 accctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct 1508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCT 4728

Qy    1509 caatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt 1568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4729 CAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGT 4788

Qy    1569 gtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgag 1628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4789 GTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAG 4848

Qy    1629 tccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc 1688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4849 TCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC 4908

Qy    1689 agagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac 1748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4909 AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC 4968

Qy    1749 actagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga 1808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4969 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5028

Qy    1809 gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgc 1868
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5029 GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGC 5088

Qy    1869 aagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg 1928
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5089 AAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG 5148

Qy    1929 gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatac 1988
           ||||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db    5149 GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCA 5208

Qy    1989 atatttgaatgtatttaga 2007
           | |      |  |  ||||
Db    5209 AAAAGGATCTTCACCTAGA 5227


Search completed: November  3, 2001, 20:34:45
Job time: 17121 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:21:54 ; Search time 5634.94 Seconds
(without alignments)
13613.261 Million cell updates/sec

Title: US-09-778-516-1
Perfect score: 8115
Sequence: 1 gatgtacgggccagatatac.....gcgtgtcgctagcgacggtc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	775	9.6	954	105	AL044364	AL044364 DKFZp434C
2	772.6	9.5	1067	107	AU081137	AU081137 AU081137
3	741.2	9.1	1089	107	AU081124	AU081124 AU081124
4	708	8.7	800	104	AJ281449	AJ281449 4A3A-P4D5
5	700.6	8.6	1004	104	AJ281480	AJ281480 4A3A-P4G8
6	677.4	8.3	700	104	AJ281616	AJ281616 4A3A-P8A1
7	672.6	8.3	1070	104	AJ281552	AJ281552 4A3A-P6F1
8	663	8.2	703	104	AJ281437	AJ281437 4A3A-P4C3
9	653.2	8.0	757	170	BF863156	BF863156 963041F03
10	621.4	7.7	734	105	AL039459	AL039459 DKFZp434O
11	615.2	7.6	718	219	AG010489	AG010489 Homo sapi
12	606.8	7.5	687	106	AU001481	AU001481 AU001481
13	604.2	7.4	1163	107	AU081044	AU081044 AU081044
14	604	7.4	636	104	AJ281699	AJ281699 4A3A-P9E3
c 15	602.6	7.4	784	235	AQ876119	AQ876119 V133E1 mT
c 16	600	7.4	795	235	AQ876011	AQ876011 V132B5 mT
c 17	599.6	7.4	756	139	BE749097	BE749097 601123138
c 18	598	7.4	959	139	BE749147	BE749147 601123315
c 19	595.4	7.3	759	139	BE749118	BE749118 601123194
c 20	594	7.3	840	139	BE749178	BE749178 601123444
21	590.4	7.3	630	105	AL042334	AL042334 DKFZp434O
22	587.2	7.2	615	105	AL044413	AL044413 DKFZp434E
23	584.8	7.2	617	104	AJ281661	AJ281661 4A3A-P8G1
24	582.8	7.2	663	106	AU001472	AU001472 AU001472
c 25	580.4	7.2	794	235	AQ875988	AQ875988 V131H3 mT
26	576.6	7.1	748	16	AI124281	AI124281 1.53 Dros
27	575.2	7.1	594	219	CI7B6	AJ227421 Ciona int
28	574.2	7.1	579	105	AL043613	AL043613 DKFZp434H
c 29	571	7.0	795	235	AQ875935	AQ875935 V131C11 m
30	568.2	7.0	628	148	BF381364	BF381364 AsIR0004
31	566.2	7.0	571	105	AL044178	AL044178 DKFZp434P
c 32	564.4	7.0	706	165	BE268532	BE268532 601125429
c 33	563.6	6.9	1045	151	BF664532	BF664532 602144525
34	563.4	6.9	690	219	AG009464	AG009464 Homo sapi
35	561.2	6.9	568	104	AJ281376	AJ281376 4A3A-P2G2
c 36	559.8	6.9	982	139	BE749172	BE749172 601123369
c 37	558.2	6.9	884	229	AQ491919	AQ491919 M902H11 m
38	558	6.9	579	104	AJ281320	AJ281320 4A3A-P1H1
c 39	555.6	6.8	857	235	AQ875006	AQ875006 V120B11 m
40	555.4	6.8	703	219	AG001761	AG001761 Homo sapi
41	553.8	6.8	615	110	AV735664	AV735664
42	553.6	6.8	616	110	AV735756	AV735756 AV735756
c 43	547.6	6.7	852	229	AQ491918	AQ491918 N900E9 mT
c 44	543.6	6.7	747	139	BE749123	BE749123 601123218
45	543.2	6.7	548	104	AJ281654	AJ281654 4A3A-P8F1

ALIGNMENTS

RESULT 1
AL044364
LOCUS AL044364 954 bp mRNA EST 29-FEB-2000
DEFINITION DKFZp434C172_s1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434C172 3', mRNA sequence.
ACCESSION AL044364
VERSION AL044364.1 GI:5432586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Ansorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansorge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
No r1 sequence available.
This clone (DKFZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES Location/Qualifiers
source 1. .954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN

Query Match 9.6%; Score 775; DB 105; Length 954;
Best Local Similarity 96.4%; Pred. No. 2.4e-205;
Matches 793; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```
Qy   1031 atgctggggatgcggtgggctctatggcttctgaggcggaaagaaccagctgcattaatg 1090
          ||| ||   | |     | | ||  | ||||   ||| |  | |   ||||||||||
Db    132 ATGGTGTCAACGTAAATGCCGCTTCGCCTTCGCGCGCGAATTGCAAGCTCTGCATTAATG 191

Qy   1091 aatcggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgct 1150
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT 251

Qy   1151 cactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggc 1210
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC 311

Qy   1211 ggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaagg 1270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 GGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGG 371

Qy   1271 ccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccg 1330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCG 431

Qy   1331 ccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacagg 1390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 CCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGG 491

Qy   1391 actataaagataccaggcgtttcccccctggaagctcccctcgtgcgctctcctgttccgac 1450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 ACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC 551

Qy   1451 cctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctca 1510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCA 611

Qy   1511 atgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgt 1570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 TAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGT 671

Qy   1571 gcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtc 1630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 GCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC 731

Qy   1631 caacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcag 1690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    732 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAG 791

Qy   1691 agcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacac 1750
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 AGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC 851

Qy   1751 tagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagt 1810
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGT 911

Qy   1811 tggtagctcttgatccggcaaacaaaccaccgctggtagcggt 1853
          |||||||||||||||||||||||||||||||||||||||||||
Db    912 TGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT 954


RESULT  2
AU081137
LOCUS       AU081137     1067 bp    mRNA            EST       15-NOV-1999
DEFINITION  AU081137 Oncorhynchus mykiss Kidney infected by infectious
            hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KI2,
            mRNA sequence.
ACCESSION   AU081137
VERSION     AU081137.1  GI:6431485
KEYWORDS    EST.
SOURCE      rainbow trout.
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1067)
  AUTHORS   Sakai,M. and Kono,T.
  TITLE     The EST analysis of kidney and gill cells rainbow trout,
            Oncorhynchus mykiss, infected with infectious hematopoietic
            necrosis virus
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Masahiro Sakai
            Faculty of Agriculture
            Miyazaki University
            1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
            Email: a0b208u@cc.miyazaki-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .1067
                     /organism="Oncorhynchus mykiss"
                     /db_xref="taxon:8022"
                     /clone="KI2"
                     /clone_lib="Oncorhynchus mykiss Kidney infected by
                     infectious hematopoietic necrosis virus"
                     /tissue_type="Kidney infected by infectious hematopoietic
                     necrosis virus"
BASE COUNT      235 a    292 c    288 g    250 t      2 others
ORIGIN


  Query Match             9.5%;  Score 772.6;  DB 107;  Length 1067;
  Best Local Similarity   98.6%;  Pred. No. 1.2e-204;
  Matches  800;  Conservative    0;  Mismatches    9;  Indels    2;  Gaps    2;

Qy   1076 ccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctc 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 CCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC 318

Qy   1136 ttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatc 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC 378

Qy   1196 agctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaa 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA 438

Qy   1256 catgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtt 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    439 CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTV 498

Qy   1316 tttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtg 1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 VTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG 558

Qy   1376 gcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcg 1435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    559 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 618

Qy   1436 ctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaag 1495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    619 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 678

Qy   1496 cgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctc 1555
          |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    679 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 738

Qy   1556 caagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaa 1615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    739 CAAGCTGGGCTGTGTGCACGAACCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA 798

Qy   1616 ctatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg 1675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 858

Qy   1676 taacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcc 1735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCC 918

Qy   1736 taactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttac 1795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 TAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC 978

Qy   1796 cttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtgg 1855
          |||||||||||||||||||||||||||||||||||||||| || ||||||||||
Db    979 CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAA-CAACGCTGGTAGCGTGTT 1037

Qy   1856 ttttttttgtttgcaagcagcagattacgcgc 1886
          |||||||||||| ||||||||||||||||||
Db   1038 TTTTTTTGTTTGC-AGCAGCAGATTACGCGC 1067


RESULT  3
AU081124
LOCUS       AU081124     1089 bp    mRNA            EST       15-NOV-1999
DEFINITION  AU081124 Oncorhynchus mykiss Kidney infected by infectious
            hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG'12,
            mRNA sequence.
ACCESSION   AU081124
VERSION     AU081124.1  GI:6431472
KEYWORDS    EST.
SOURCE      rainbow trout.
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1089)
  AUTHORS   Sakai,M. and Kono,T.
  TITLE     The EST analysis of kidney and gill cells rainbow trout,
            Oncorhynchus mykiss, infected with infectious hematopoietic
            necrosis virus
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Masahiro Sakai
            Faculty of Agriculture
            Miyazaki University
            1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
            Email: a0b208u@cc.miyazaki-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .1089
                     /organism="Oncorhynchus mykiss"
                     /db_xref="taxon:8022"
                     /clone="KG'12"
                     /clone_lib="Oncorhynchus mykiss Kidney infected by
                     infectious hematopoietic necrosis virus"
```

```
                /tissue_type="Kidney infected by infectious hematopoietic
                necrosis virus"
BASE COUNT      245 a    297 c    295 g    251 t      1 others
ORIGIN


  Query Match             9.1%;  Score 741.2;  DB 107;  Length 1089;
  Best Local Similarity   96.5%;  Pred. No. 7.5e-196;
  Matches  801;  Conservative    0;  Mismatches   23;  Indels    6;  Gaps    4;

Qy   1076 ccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctc 1135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    260 CCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC 319

Qy   1136 ttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatc 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    320 TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC 379

Qy   1196 agctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaa 1255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    380 AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA 439

Qy   1256 catgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtt 1315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    440 CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT 499

Qy   1316 tttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtg 1375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    500 TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG 559

Qy   1376 gcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcg 1435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    560 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 619

Qy   1436 ctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaag 1495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    620 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 679

Qy   1496 cgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctc 1555
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    680 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 739

Qy   1556 caagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaa 1615
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    740 CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA 799

Qy   1616 ctatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg 1675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    800 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 859

Qy   1676 taacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaag-tggtggc 1734
           |||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db    860 TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTTTGTGGC 919

Qy   1735 ctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagtta 1794
           |||||||||||||||||| |||||||| |||||||||||||||| |||||||||||||||
Db    920 CTAACTACGGCTACACTAGAAGAACAGTATTTTGTATCTGCGCTCTGCTGGAGCCAGTTA 979

Qy   1795 ccttcggaaaaagagttggtagctctt-gatccggcaaacaaacca---ccgctggtagc 1850
           |||||||||||||||| |||||||| |||||||||||||||| ||    |   |   |||
Db    980 CCTTCGGAAAAAGAGTGGGTAGCTCTTGGATCCGGCAAACAAAACAACCGCTGGGAAGGG 1039

Qy   1851 ggtggtttttttgtttgcaagca-gcagattacgcgcagaaaaaaaggat 1899
           ||  || ||||||||||||||| |||||||| |||  |||||| ||||
Db   1040 GGGTGTCTTTTTGTTTGCAAGCACGCAGATTACCCGCGAAAAAAACGGAT 1089


RESULT   4
AJ281449
LOCUS       AJ281449      800 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION   AJ281449
VERSION     AJ281449.1  GI:6929329
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .800
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P4D5"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT      203 a    206 c    198 g    193 t
ORIGIN


  Query Match             8.7%;  Score 708;  DB 104;  Length 800;
  Best Local Similarity   96.7%;  Pred. No. 1.3e-186;
  Matches  723;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy   1260 tgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttc 1319
           | |||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 TAAGCTTGCGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC 71

Qy   1320 cataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcga 1379
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     72 CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATAGACGCTCAAGTCAGAGGTGGCGA 131

Qy   1380 aacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctct 1439
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    132 AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT 191

Qy   1440 cctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtg 1499
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG 251

Qy   1500 gcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaag 1559
           |||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db    252 GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG 311

Qy   1560 ctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactat 1619
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT 371

Qy   1620 cgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaac 1679
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC 431
```

```
Qy   1680 aggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaac 1739
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 491

Qy   1740 tacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttc 1799
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    492 TACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC 551

Qy   1800 ggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttt 1859
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT 611

Qy   1860 tttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatc 1919
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC 671

Qy   1920 ttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatg 1979
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG 731

Qy   1980 agcggatacatatttgaatgtatttaga 2007
          ||  ||  | |      |  |  ||||
Db    732 AGATTATCAAAAAGGATCTTCACCTAGA 759


RESULT   5
AJ281480
LOCUS       AJ281480     1004 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION   AJ281480
VERSION     AJ281480.1  GI:6929360
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 1004)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .1004
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P4G8"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT      252 a    262 c    244 g    244 t      2 others
ORIGIN


  Query Match             8.6%;  Score 700.6;  DB 104;  Length 1004;
  Best Local Similarity   97.3%;  Pred. No. 1.7e-184;
  Matches  712;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1276 aaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccc 1335
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC 60

Qy   1336 ctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactat 1395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTAT 120

Qy   1396 aaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgc 1455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC 180

Qy   1456 cgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgct 1515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db    181 CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCT 240

Qy   1516 cacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacg 1575
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    241 CACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTNTGCACG 300

Qy   1576 aacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacc 1635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACC 360

Qy   1636 cggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcga 1695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA 420

Qy   1696 ggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaa 1755
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAA 480

Qy   1756 ggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggta 1815
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTA 540

Qy   1816 gctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagc 1875
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGC 600

Qy   1876 agattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctg 1935
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG 660

Qy   1936 acgctcagtggaacgaaaactcacgttaagggattttggtcatgagcggatacatatttg 1995
          ||||||||||||||||||||||||||||||||||||||||||||||||||  ||  | |
Db    661 ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGA 720

Qy   1996 aatgtatttaga 2007
            |  |  ||||
Db    721 TCTTCACCTAGA 732


RESULT   6
AJ281616
LOCUS       AJ281616     700 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P8A12-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P8A12, mRNA sequence.
ACCESSION   AJ281616
VERSION     AJ281616.1  GI:6929495
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
```

```
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P8A12"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT      167 a    196 c    189 g    146 t      2 others
ORIGIN


  Query Match             8.3%;  Score 677.4;  DB 104;  Length 700;
  Best Local Similarity   98.8%;  Pred. No. 4.8e-178;
  Matches  681;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1163 gcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggtt 1222
           ||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12 GCTTGCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTT 71

Qy   1223 atccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggc 1282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     72 ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC 131

Qy   1283 caggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccctgacga 1342
           |||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db    132 CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGNNTTTCCATAGGCTCCGCCCCCCTGACGA 191

Qy   1343 gcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagata 1402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    192 GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA 251

Qy   1403 ccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttac 1462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    252 CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC 311

Qy   1463 cggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctg 1522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    312 CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG 371

Qy   1523 taggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccc 1582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    372 TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC 431

Qy   1583 cgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaag 1642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAG 491

Qy   1643 acacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgt 1702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT 551

Qy   1703 aggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagt 1762
           |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    552 AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGT 611

Qy   1763 atttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttg 1822
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG 671

Qy   1823 atccggcaaacaaaccaccgctggtagcg 1851
           |||||||||||||||||||||||||||||
Db    672 ATCCGGCAAACAAACCACCGCTGGTAGCG 700


RESULT   7
AJ281552
LOCUS       AJ281552     1070 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION   AJ281552
VERSION     AJ281552.1  GI:6929432
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 1070)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .1070
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P6F11"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT      263 a    283 c    255 g    269 t
ORIGIN


  Query Match             8.3%;  Score 672.6;  DB 104;  Length 1070;
  Best Local Similarity   97.3%;  Pred. No. 1.2e-176;
  Matches  684;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   1305 ttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctca 1364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCA 60

Qy   1365 agtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagc 1424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGC 120

Qy   1425 tccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctc 1484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTC 180
```

```
Qy   1485 ccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtag 1544
          ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    181 CCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG 240

Qy   1545 gtcgttcgctccaagctgggctgtgtgcacgaacccccccgttcagcccgaccgctgcgcc 1604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCC 300

Qy   1605 ttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggca 1664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA 360

Qy   1665 gcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttg 1724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG 420

Qy   1725 aagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctg 1784
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    421 AAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG 480

Qy   1785 aagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgct 1844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT 540

Qy   1845 ggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaa 1904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAA 600

Qy   1905 gaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaa 1964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAA 660

Qy   1965 gggattttggtcatgagcggatacatatttgaatgtatttaga 2007
          |||||||||||||||   ||  | |      |  |  ||||
Db    661 GGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGA 703
```

```
RESULT   8
AJ281437
LOCUS       AJ281437      703 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P4C3-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P4C3, mRNA sequence.
ACCESSION   AJ281437
VERSION     AJ281437.1  GI:6929317
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P4C3"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT      159 a    204 c    193 g    147 t
ORIGIN
```

```
  Query Match             8.2%;  Score 663;  DB 104;  Length 703;
  Best Local Similarity   97.8%;  Pred. No. 5.2e-174;
  Matches  672;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1075 accagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgct 1134
          || |   | | |||     ||||||||||||||||||||||||||||||||||||||||||||
Db     17 ACTAAAGGGAATAAGCTTGCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT 76

Qy   1135 cttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtat 1194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTAT 136

Qy   1195 cagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaaga 1254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA 196

Qy   1255 acatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgt 1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT 256

Qy   1315 ttttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggt 1374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 TTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT 316

Qy   1375 ggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgc 1434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC 376

Qy   1435 gctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaa 1494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA 436

Qy   1495 gcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgct 1554
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    437 GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT 496

Qy   1555 ccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttatccggta 1614
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTA 556

Qy   1615 actatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactg 1674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 ACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG 616

Qy   1675 gtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggc 1734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    617 GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGC 676

Qy   1735 ctaactacggctacactagaaggacag 1761
          ||||||| ||||||||||||||| ||||
Db    677 CTAACTAGGGCTACACTAGAAGAACAG 703
```

```
RESULT   9
BF863156
LOCUS       BF863156      757 bp    mRNA            EST       19-JAN-2001
DEFINITION  963041F03.x1 C. reinhardtii CC-1690, Stress condition I, normalized
            , Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION   BF863156
VERSION     BF863156.1  GI:12253300
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
```

```
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C.,
            Lefebvre,P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
  TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants; project phase 3
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Duke University
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1. .757
                     /organism="Chlamydomonas reinhardtii"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Stress condition I,
                     normalized, Lambda Zap II"
                     /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI; This library, constructed by John Davies and Jeffrey
                     McDermott, combines cDNAs from CC-1690 cells grown to
                     mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
                     1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr
                     , 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
                     purified from each sample, pooled and cDNA synthesized.
                     The cDNA was directionally cloned into lambda Zap II
                     (Stratagene) in the EcoRI (5') and XhoRI (3') sites.
                     pBluescript II SK- plasmids were excised from the lambda
                     ZAP clones by superinfection with ExAssist (Stratagene)
                     phage. The library was normalized using method 4 described
                     in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT      167 a    219 c    208 g    161 t      2 others
ORIGIN
```

```
  Query Match             8.0%;  Score 653.2;  DB 170;  Length 757;
  Best Local Similarity   95.3%;  Pred. No. 3e-171;
  Matches  684;  Conservative    0;  Mismatches   33;  Indels    1;  Gaps    1;

Qy   1035 tggggatgcggtgggctctatggcttctgaggcggaaagaaccagctgcattaatgaatc 1094
          |||| |   ||    | || |   |   ||  |     |  ||||||||||||||||||
Db     34 TGGGTACCGGGCCCCCGCTTTCCGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC 93

Qy   1095 ggccaacgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcact 1154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACT 153

Qy   1155 gactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggta 1214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 GACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTA 213

Qy   1215 atacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccag 1274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 ATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAG 273

Qy   1275 caaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccc 1334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 333

Qy   1335 cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggacta 1394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA 393

Qy   1395 taaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctg 1454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTG 453
```

```
Qy   1455 ccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgc 1514
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db    454 CCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC 513

Qy   1515 tcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcac 1574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC 573

Qy   1575 gaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaac 1634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GAACCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAAC 633

Qy   1635 ccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcg 1694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 CCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCG 693

Qy   1695 aggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacacta 1752
          |||||||||||||||||||||||||||||| ||||     ||||| ||||||||||||
Db    694 AGGTATGTAGGCGGTGCTACAGAGTTCTTGGAGTGAGTAGCTAAC-ACGGCTACACTA 750
```

```
RESULT  10
AL039459
LOCUS       AL039459      734 bp    mRNA            EST       29-FEB-2000
DEFINITION  DKFZp434O1210_s1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434O1210 3', mRNA sequence.
ACCESSION   AL039459
VERSION     AL039459.2  GI:5935158
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 734)
  AUTHORS   Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
            ,S.
  TITLE     EST (Duesterhoeft, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     On Jul 7, 1999 this sequence version replaced gi:5408507.
            Contact: Duesterhoeft A
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No r1 sequence available.
            This clone (DKFZp434O1210) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .734
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434O1210"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      172 a    199 c    194 g    161 t      8 others
ORIGIN
```

```
  Query Match             7.7%;  Score 621.4;  DB 105;  Length 734;
  Best Local Similarity   94.4%;  Pred. No. 2.4e-162;
  Matches  685;  Conservative    0;  Mismatches   33;  Indels    8;  Gaps    4;

Qy   1159 cgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatac 1218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      9 CGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATAC 68
```

```
Qy   1219 ggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaa 1278
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     69 GGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAA 128

Qy   1279 aggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctg 1338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 AGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCCTG 188

Qy   1339 acgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaa 1398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA 248

Qy   1399 gataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgc 1458
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    249 GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGC 308

Qy   1459 ttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcac 1518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    309 TTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC 368

Qy   1519 gctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaac 1578
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    369 GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC 428

Qy   1579 cccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaaccc-g 1637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    429 CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG 488

Qy   1638 gtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgagg 1697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    489 GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG 548

Qy   1698 tatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagg 1757
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    549 TATGTAGGCGGTGCTACANAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGA 608

Qy   1758 acagta-tttggtatctgcgctctgctgaagccagttaccttcggaaaaa--gagttggt 1814
          |||||| |||||||||||||||||||||| ||||||||||||||||||    || |||
Db    609 ACAGTATTTTGGTATCTGCGCTCTGCTGAAACCAGTTACCTTCGGAAAAAAGANGTGGGT 668

Qy   1815 agctcttg----atccggcaaacaaaccaccgctggtagcggtggttttttttgtttgcaa 1870
          ||| ||||        |   ||| || | |||||||||     |  | ||||||| | |
Db    669 AGCNCTTGGATTCGGCNAAAAAAAACCAACCGCTGGTTAGNGGTGGTTTTTTTTNTTTNC 728

Qy   1871 gcagca 1876
          ||||
Db    729 AAAGCA 734


RESULT  11
AG010489
LOCUS       AG010489      718 bp    DNA             GSS       14-APR-1999
DEFINITION  Homo sapiens genomic DNA, 21q region, clone: P7H8X25, genomic
            survey sequence.
ACCESSION   AG010489 AG004363
VERSION     AG010489.1  GI:3327308
KEYWORDS    GSS.
SOURCE      Homo sapiens DNA, clone:P7H8X25.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA, chromosome 21q
  JOURNAL   Published Only in DataBase (1998) In press
REFERENCE   2  (bases 1 to 718)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-1998) to the DDBJ/EMBL/GenBank databases.
            Masahira Hattori, Kitasato University, Department of Science, JST
            Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan
            (E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
            Fax:0427-78-9561)
COMMENT     On Feb 5, 1999 this sequence version replaced gi:2789511.
            AG004363: Submitted (17-Jan-1998).
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /clone="P7H8X25"
                     /map="21q"
BASE COUNT      176 a    191 c    187 g    158 t      6 others
ORIGIN


  Query Match             7.6%;  Score 615.2;  DB 219;  Length 718;
  Best Local Similarity   97.2%;  Pred. No. 1.3e-160;
  Matches  667;  Conservative    0;  Mismatches   15;  Indels    4;  Gaps    4;

Qy   1224 tccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcc 1283
          ||| | | |  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 TCCCCGGGAGGAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC 80

Qy   1284 aggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgacgag 1343
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAG 140

Qy   1344 catcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagatac 1403
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 CATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC 200

Qy   1404 caggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttacc 1463
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACC 260

Qy   1464 ggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgt 1523
          |||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db    261 GGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGT 320

Qy   1524 aggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccc 1583
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    321 AGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC 380

Qy   1584 gttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaaga 1643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    381 GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGA 440

Qy   1644 cacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgta 1703
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db    441 CACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCNAGGTATGTA 500

Qy   1704 ggcggtgct-acagagttcttgaagtggtggcctaactacggct-acactagaaggacag 1761
          ||||||||| |||||||||||||||||||||||||||||||| |||||||||| ||||
Db    501 GGCGGTGCTAACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAACACTAGAAGAACAG 560

Qy   1762 tatttggtatctgcgctctgctgaagccagttacc-ttcggaaaaagagttggtagctct 1820
          |||||||||||||||||||||||||||||||| | ||||||||||||||||||||||||
Db    561 TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTANCTTTCGGAAAAAGAGTTGGTAGCTCT 620

Qy   1821 tgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagcagcagatt 1880
          ||||| |||||||||||||||||||| ||| ||||||  ||||||||||||||||
Db    621 TGATCGGGCAAACAAACCACCGCTGGTAGC-GTGTTTTTTTGTTTTGCAAGCAGCAGATA 679

Qy   1881 acgcgcagaaaaaaaaggatctcaaga 1906
          |||||||||||||||||||||||||||
Db    680 ACGCGCAGAAAAAAAGGATCTCAAGA 705


RESULT  12
AU001481
LOCUS       AU001481      687 bp    mRNA            EST       15-JAN-1999
DEFINITION  AU001481 Bombyx mori p50(Daizo) Bombyx mori cDNA clone fbm0478f,
            mRNA sequence.
ACCESSION   AU001481
VERSION     AU001481.1  GI:4157725
```

```
KEYWORDS    EST.
SOURCE      domestic silkworm.
  ORGANISM  Bombyx mori
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
            ; Bombycoidea; Bombycidae; Bombyx.
REFERENCE   1  (bases 1 to 687)
  AUTHORS   Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
  TITLE     Establishment of cDNA database of Bombyx mori
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Mita K
            Genome Research Group
            National Institute of Radiological Sciences
            Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
            Email: kmita@nirs.go.jp
            PROJECT = 'CREST project by JST'.
FEATURES             Location/Qualifiers
     source          1. .687
                     /organism="Bombyx mori"
                     /strain="p50(Daizo)"
                     /db_xref="taxon:7091"
                     /clone="fbm0478f"
                     /clone_lib="Bombyx mori p50(Daizo)"
BASE COUNT      145 a    209 c    183 g    149 t      1 others
ORIGIN
```

```
  Query Match             7.5%;  Score 606.8;  DB 106;  Length 687;
  Best Local Similarity   99.5%;  Pred. No. 2.9e-158;
  Matches  608;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy    1076 ccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctc 1135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 CCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC 136

Qy    1136 ttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatc 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC 196

Qy    1196 agctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaa 1255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA 256

Qy    1256 catgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtt 1315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTN 316

Qy    1316 tttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtg 1375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG 376

Qy    1376 gcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcg 1435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     377 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 436

Qy    1436 ctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaag 1495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     437 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 496

Qy    1496 cgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctc 1555
           ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db     497 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 556

Qy    1556 caagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaa 1615
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     557 CAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAA 616

Qy    1616 ctatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg 1675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     617 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 676

Qy    1676 taacaggatta 1686
           ||||||||||||
Db     677 TAACAGGATTA 687
```

```
RESULT  13
AU081044
LOCUS       AU081044     1163 bp    mRNA            EST       15-NOV-1999
DEFINITION  AU081044 Oncorhynchus mykiss Kidney infected by infectious
            hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KB1,
            mRNA sequence.
ACCESSION   AU081044
VERSION     AU081044.1  GI:6431392
KEYWORDS    EST.
SOURCE      rainbow trout.
  ORGANISM  Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1  (bases 1 to 1163)
  AUTHORS   Sakai,M. and Kono,T.
  TITLE     The EST analysis of kidney and gill cells rainbow trout,
            Oncorhynchus mykiss, infected with infectious hematopoietic
            necrosis virus
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Masahiro Sakai
            Faculty of Agriculture
            Miyazaki University
            1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
            Email: a0b208u@cc.miyazaki-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .1163
                     /organism="Oncorhynchus mykiss"
                     /db_xref="taxon:8022"
                     /clone="KB1"
                     /clone_lib="Oncorhynchus mykiss Kidney infected by
                     infectious hematopoietic necrosis virus"
                     /tissue_type="Kidney infected by infectious hematopoietic
                     necrosis virus"
BASE COUNT      272 a    307 c    298 g    286 t
ORIGIN
```

```
  Query Match             7.4%;  Score 604.2;  DB 107;  Length 1163;
  Best Local Similarity   99.4%;  Pred. No. 1.9e-157;
  Matches  617;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy    1076 ccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgctc 1135
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     544 CCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTC 603

Qy    1136 ttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatc 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     604 TTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATC 663

Qy    1196 agctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaa 1255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     664 AGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAA 723

Qy    1256 catgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtt 1315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     724 CATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT 783

Qy    1316 tttccataggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtg 1375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     784 TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG 843

Qy    1376 gcgaaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcg 1435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     844 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCG 903

Qy    1436 ctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaag 1495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     904 CTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAG 963
```

```
Qy   1496 cgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctc 1555
          ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db    964 CGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTC 1023

Qy   1556 caagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaa 1615
          |||||||||||||||||||| |||||||||||||||||||| ||||||||||
Db   1024 CAAGCTGGGCTGTGTGCACGAA-CCCCCGTTCAGCCCGACCGCTGCGCTTTATCCGGTAA 1082

Qy   1616 ctatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg 1675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1083 CTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG 1142

Qy   1676 taacaggattagcagagcgag 1696
          ||||||||||||||||||||
Db   1143 TAACAGGATTAGCAGAGCGAG 1163


RESULT 14
AJ281699
LOCUS       AJ281699      636 bp    mRNA            EST       30-JUN-2000
DEFINITION  4A3A-P9E3-F Anopheles gambiae immune competent 4A3A Anopheles
            gambiae cDNA clone 4A3A-P9E3, mRNA sequence.
ACCESSION   AJ281699
VERSION     AJ281699.1  GI:6929578
KEYWORDS    EST.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
            ; Anopheles.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
            Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
            and Kafatos,F.C.
  TITLE     Anopheles gambiae pilot gene discovery project: identification of
            mosquito innate immunity genes from expressed sequence tags
            generated from immune-competent cell lines
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
  MEDLINE   20300950
COMMENT     Contact: Dimopoulos G
            Fotis C. Kafatos laboratory
            European Molecular Biology Laboratory
            Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P9E3"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT      152 a    179 c    166 g    139 t
ORIGIN


  Query Match             7.4%;  Score 604;  DB 104;  Length 636;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-157;
  Matches  629;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    2;

Qy   1202 ctcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtg 1261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG 60

Qy   1262 agcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttcca 1321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA 120

Qy   1322 taggctccgccccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaa 1381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAA 180

Qy   1382 cccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcc 1441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC 240

Qy   1442 tgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggc 1501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC 300

Qy   1502 gctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagct 1561
          |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT 360

Qy   1562 gggctgtgtgcacgaacccccccgttcagcccgaccgctgcgccttatccggtaactatcg 1621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCG 420

Qy   1622 tcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacag 1681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG 480

Qy   1682 gattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcc-taact 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |  ||
Db    481 GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTTTTCT 540

Qy   1741 acggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcg 1800
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    541 ACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 600

Qy   1801 g-aaaaagagttggtagctcttgatccggcaaacaa 1835
          | ||||||||||||||||||||||||||||||||||
Db    601 GAAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA 636


RESULT 15
AQ876119/c
LOCUS       AQ876119      784 bp    DNA             GSS       08-NOV-1999
DEFINITION  V133E1 mTn-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
            cerevisiae genomic 5', DNA sequence.
ACCESSION   AQ876119
VERSION     AQ876119.1  GI:6288363
KEYWORDS    GSS.
SOURCE      Saccharomyces cerevisiae.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A.,
            desEtages,S.A., Cheung,K.-H., Sheehan,A., Symoniatis,D., Jansen,R.,
            Umansky,L., Heidtman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,
            Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
  TITLE     Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
            Gene Disruption
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Kumar A
            Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
            Yale University
            P.O. Box 208103, New Haven, CT 06520-8103, USA
            Tel: 203 432 9949
            Fax: 203 432 6161
            Email: anuj.kumar@yale.edu
            te of mTn-3xHA/lacZ insertion.
            Seq primer: GGCCTTCTTTCTTTGGAAGTAC
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Saccharomyces cerevisiae"
```

```
                     /strain="Y2278 - S288C background, cir(0) rho(0)"
                     /db_xref="taxon:4932"
                     /clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
                     /lab_host="E. coli"
                     /note="Vector: pHSS6-Sal; A yeast genomic DNA library
                     without 2 micron or mitochondrial DNA  was prepared in
                     pHSS6-Sal; genomic DNA was size-fractionated  (DNA of
                     roughly 2-3 kb in length) prior to cloning. This library
                     was subsequently mutagenized with a mTn-3xHA/lacZ
                     minitransposon  containing lacZ, URA3, and tet resistance.
                     "
BASE COUNT      187 a    200 c    210 g    186 t      1 others
ORIGIN


  Query Match             7.4%;  Score 602.6;  DB 235;  Length 784;
  Best Local Similarity   97.6%;  Pred. No. 4.5e-157;
  Matches  611;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy   1304 gttgctggcgtttttccataggctccgccccccctgacgagcatcacaaaaatcgacgctc 1363
           | ||||||||||||||||||||||| |||||| ||||||||||||||||||||||||||
Db    784 GGTGCTGGCGTTTTTCCATAGGCTCGGCCCCCTTGACGAGCATCACAAAAATCGACGCTC 725

Qy   1364 aagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccctggaag 1423
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    724 AAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATCCCAGGCGTTTCCCCCTGGAAG 665

Qy   1424 ctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttct 1483
           |||||||||||||||||||||||||| ||| |||||||| ||||||||| |||| |||
Db    664 CTCCCTCGTGCGCTCTCCTGTTCCGACCTTGCTGCTTACCGAATACCTGTCAGCCTGTCT 605

Qy   1484 cccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgta 1543
           |||||||| |||||||||||||| |||  ||||||||||||||||||||||||||||||
Db    604 CCCTTCGGNAAGCGTGGCGCTTTATCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTA 545

Qy   1544 ggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgc 1603
           |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    544 GGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCTCCGTTCAGCCCGACCGCTGCGC 485

Qy   1604 cttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggc 1663
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 CTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC 425

Qy   1664 agcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttctt 1723
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT 365

Qy   1724 gaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgct 1783
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCT 305

Qy   1784 gaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgc 1843
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 GAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC 245

Qy   1844 tggtagcggtggttttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctca 1903
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 TGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA 185

Qy   1904 agaagatcctttgatcttttctacgg 1929
           ||||||||||||||||||||||| |
Db    184 AGAAGATCCTTTGATCTTTTCTACTG 159


Search completed: November  3, 2001, 17:35:03
Job time: 7989 sec
```